LES DOSSIERS DE l'iecn



# LCA

## COLLÈGE UNIVERSITAIRE DES ENSEIGNANTS DE SANTÉ PUBLIQUE

- Dossiers format iECN (QRM, QRU)
- Commentaires et rappels de cours de spécialistes
- En accord avec les recommandations et conférences de consensus

Articles à télécharger

**CUEP**
Collège Universitaire des Enseignants de Santé Publique

ellipses

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# LCA

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

ISBN : 9782340-025868
©Ellipses Édition Marketing S.A., 2018
32, rue Bargue 75740 Paris cedex 15




Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5.2° et 3°a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective », et d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).
Cette représentation ou reproduction, par quelque procédé que ce soit constituerait une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

www.editions-ellipses.fr

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Préface

Écrire une préface à un ouvrage d'apprentissage et d'entraînement à la lecture critique d'article est une gageure. Les champs méthodologiques couverts sont très vastes, les thèmes de recherche sont infinis et la capacité d'un chercheur à transmettre fidèlement le travail de son équipe dans les conditions fixées par la revue scientifique conduisent à des formats de lecture peu aisés.

Dès les années 1980, un certain nombre de nos collègues avaient pensé à enseigner la Lecture Critique d'Articles médicaux aux étudiants en Médecine. L'idée première était de les former à acquérir une vision critique de la littérature scientifique et de les aider à tester leurs capacités de synthèse par l'élaboration d'un résumé.

Cette épreuve, d'abord sous forme papier, s'est mise en place progressivement il y a dix ans, et aujourd'hui, informatisée, elle compte pour 10 % de la note finale dans les Épreuves Classantes Nationales Informatisées. Chaque épreuve de Lecture Critique est constituée de deux articles originaux de bonne qualité en Anglais, avec 15 Questions à Réponse Multiple construites de manière progressive (de 13 à 17 selon les textes).

L'objectif de la Lecture Critique d'Articles est l'acquisition pour les futurs médecins d'une compétence indispensable pour leur formation continue : être capable d'analyser le rapport d'une étude afin de décider si les résultats doivent conduire à un changement de ses pratiques.

Ce qui est testé dans ce type d'épreuve est la capacité à comprendre tous les éléments de l'étude pour avoir une vision claire des informations qui peuvent en être tirées. L'une des bases du développement professionnel continu est de se tenir informé tout au long de sa carrière des avancées scientifiques. L'apprentissage de la Lecture Critique fournit les éléments de cette formation que les jeunes médecins continueront à exercer au cours de leur internat par les réunions de bibliographie organisées dans de très nombreux services.

Pour profiter pleinement de l'ouvrage, il faut se plonger dans la lecture bien sûr, mais probablement aussi se replonger dans les fondements de l'activité scientifique telle que Thomas Kuhn l'a définie. L'activité de recherche scientifique est une révolution permanente et ordonnée de l'association construite entre la théorie, les méthodes et les faits collectés. On retrouve dans cette triade l'association entre l'objectif – l'idée de la recherche-, les méthodes épidémiologiques et statistiques employées, et les données collectées dans la population des sujets ou des patients inclus dans une étude publiée.

Il reste aussi à apprécier, et c'est une question de culture générale médicale, si tout cet assemblage d'idées, de règles méthodologiques et de conventions d'écriture, produisent un résultat applicable à la pratique clinique ou à la pratique de santé publique dans notre système de santé européen en général et français en particulier.

Ce livre a été réalisé dans cet esprit. C'est la compilation d'un certain nombre d'exemples avec des questions qui ont été revues collégialement par plusieurs groupes de professionnels de Santé Publique. Les réponses relativement détaillées qui sont données en correction des questions, permettent de mettre à jour les connaissances voire de les affiner.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Nous souhaitons que cet ouvrage puisse apporter une aide concrète à nos étudiants allant bien au-delà de la réussite à l'Examen Classant National, c'est-à-dire se familiariser à une question essentielle de la pratique médicale : suis-je en mesure de proposer à mes patients les meilleurs soins possibles au regard de la médecine basée sur les preuves ?

Pr Cyrille Colin
*Président du Collège Universitaire des Enseignants de Santé Publique*

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Table des matières

**Préface** ..... 3
**Entraînement à la LCA pour les ECN : mode d'emploi** ..... 7
**Les auteurs** ..... 9

**ESSAIS THÉRAPEUTIQUES/ ESSAIS D'INTERVENTION**

**Dossier 1** *Essai de supériorité* ..... 15
**Dossier 2** *Essai de supériorité* ..... 25
**Dossier 3** *Essai de supériorité* ..... 35
**Dossier 4** *Essai de supériorité* ..... 47
**Dossier 5** *Essai d'équivalence* ..... 59
**Dossier 6** *Essai de non-infériorité* ..... 69
**Dossier 7** *Essai de non-infériorité* ..... 79
**Dossier 8** *Essai en cross-over* ..... 91
**Dossier 9** *Essai en cross-over* ..... 99
**Dossier 10** *Essai en plan factoriel* ..... 107
**Dossier 11** *Essai d'intervention* ..... 117
**Dossier 12** *Essai d'intervention* ..... 127
**Dossier 13** *Essai d'intervention* ..... 135
**Dossier 14** *Essai de prévention* ..... 143
**Dossier 15** *Essai de prévention* ..... 151

**ÉTUDES OBSERVATIONNELLES**

**Dossier 1** *Étude de cohorte* ..... 165
**Dossier 2** *Étude de cohorte (score de propension)* ..... 171
**Dossier 3** *Étude de cohorte* ..... 181
**Dossier 4** *Étude de cohorte (cohorte historique)* ..... 193
**Dossier 5** *Étude cas-témoin* ..... 205
**Dossier 6** *Étude cas-témoin* ..... 215
**Dossier 7** *Étude cas-témoin (nichée dans une cohorte)* ..... 225
**Dossier 8** *Étude cas-témoin* ..... 235
**Dossier 9** *Étude descriptive (transversale)* ..... 243
**Dossier 10** *Étude descriptive (sur registre)* ..... 255
**Dossier 11** *Étude diagnostique* ..... 265
**Dossier 12** *Étude diagnostique* ..... 273
**Dossier 13** *Étude de pratique* ..... 283

**ANNALES /ECN BLANC**

**Dossier 1** *Épreuve de LCA 2016* ..... 293
**Dossier 2** *Épreuve blanche nationale LCA 2017* ..... 303



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Entraînement à la LCA pour les ECN : mode d'emploi

S'il existe quelques discordances entre les enseignants sur certains points de détails en rapport avec l'enseignement de la lecture critique d'article, la clé de votre réussite réside dans les principes qui mettent tout le monde d'accord :

1. à l'issue de votre deuxième cycle, vous devez être capable de lire et de comprendre l'essentiel d'un article de recherche en 30 minutes ;
2. les aspects techniques, notamment statistiques n'ont d'intérêt que s'ils sont nécessaires à l'interprétation des résultats ou au jugement de la validité de l'étude présentée ;
3. derrière chaque QCM, l'enseignant cache plus ou moins subtilement un objectif pédagogique que vous avez certainement déjà entendu ou lu quelque part.

Il est probable qu'en LCA, l'entraînement soit plus important encore que dans les autres matières. En effet, les connaissances théoriques nécessaires à l'abord de la matière peuvent s'acquérir relativement rapidement. En revanche l'aptitude à la lecture « rapide » d'un article en anglais est le savoir-faire qui fera la différence le jour J. Seul l'entraînement vous permettra d'acquérir les réflexes de lecture qui vous permettront de gagner de précieuses minutes pour répondre aux QCM de vos examens facultaires et de l'épreuve finale.

Le jour J, l'épreuve comportera deux articles de recherche en anglais, dont au moins un essai clinique. Chaque article sera accompagné de 15 questions à réponse unique ou multiple, présentées à l'étudiant sur un mode « progressif » (les QCM sont obligatoirement complétés dans l'ordre défini par l'enseignant, sans retour en arrière possible). La durée de l'épreuve est de 3 heures.

Pour être utile à votre préparation, l'abord de chaque sujet présenté dans cet ouvrage doit vous rapprocher le plus possible de ces conditions. Les auteurs vous proposent donc une approche en trois temps de chaque article, par sessions d'1 h 30 à 2 h 00 :

**1. La lecture de l'article (30 à 45 minutes maximum)**

Lors des examens facultaires, trop d'étudiants sont encore en train de détailler l'article plus d'une heure après le début de l'épreuve. À l'inverse, lors des entraînements et conférences, cette étape est souvent « bâclée » en 15 minutes. Notre expérience montre que le meilleur compromis est de passer environ 30 minutes de concentration sur l'article avant d'envisager de répondre aux questions. Peu importe le détail de l'analyse statistique ou de la description des paramètres biologiques recueillis, ce temps est nécessaire et suffisant pour 1) « habituer l'œil » à l'article et ainsi repérer où sont les éléments clés et 2) pour en comprendre l'essentiel :

a. objectifs ;
b. type d'étude ;
c. population d'étude ;
d. critère de jugement principal ;
e. nature de la maladie et de l'exposition (étude épidémiologique) ou de l'intervention (essai clinique) ;



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

f. caractéristiques de la randomisation et de la procédure d'aveugle (essais cliniques) ;
g. principaux résultats, compréhension des tableaux et figures.

En fonction de votre niveau d'anglais et de vos connaissances théoriques en épidémiologie/statistiques, il est normal de ne pas être capable de réussir cet exercice dans le temps imparti en début de DFASM2. L'entraînement à la lecture est donc la clé du succès !

**2. La réponse aux QCM (30 à 45 minutes, soit 2 à 4 par QCM)**

Après avoir lu **l'ensemble des propositions du QCM**, deux questions doivent se poser immédiatement :

1. Quel est l'objectif pédagogique du QCM ? (ex : savoir qu'un critère de jugement subjectif pose problème dans un essai ouvert, savoir que la connaissance de la maladie peut influencer la mesure rétrospective d'une exposition, comprendre pourquoi l'analyse per-protocole est plus conservatrice dans un essai de non-infériorité...). Si vous arrivez à formuler cet objectif à la lecture de l'intitulé de la question, l'ensemble des propositions doit être ensuite critiqué à la lumière du point de cours que vous avez identifié.

2. Y a-t-il une ou des propositions que je peux éliminer d'emblée ? Certaines proposition sont fausses dans l'absolu (ex : dans cette étude l'analyse en per protocole comprenait plus de patients que l'analyse en intention de traiter, tous les facteurs de confusions possibles ont été pris en compte...).

La stratégie consiste donc à éliminer, sur votre brouillon, les propositions fausses dans l'absolu, s'il y en a, puis à retourner chercher dans le texte les informations nécessaires à la réponse aux autres propositions en gardant à l'esprit l'objectif pédagogique du QCM en cas de doute.

**3. L'auto-évaluation**

Il est conseillé de ne découvrir la correction qu'après avoir répondu à l'ensemble du sujet proposé. La notation des QCM est faite de la manière suivante :
– toutes les bonnes réponses : 1 point ;
– une discordance (une réponse manquante ou une mauvaise réponse cochée) : 0,5 points ;
– deux discordances (réponses manquantes ou cochées à tort) : 0,25 points ;
– plus de deux discordances : 0 point.

La lecture des commentaires est indispensable pour comprendre les raisons de vos erreurs et revoir les points de cours correspondants.

Bonne analyse !



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Les auteurs

Nelly AGRINIER
*Maître de Conférences en épidémiologie, économie de la santé et prévention*
*UFR de Médecine, Université de Lorraine, Nancy*

Étienne AUDUREAU
*Professeur en épidémiologie, économie de la santé et prévention*
*UFR de Médecine, Université Paris-Est Créteil*

Pascal AUQUIER
*Professeur en épidémiologie, économie de la santé et prévention*
*UFR de médecine, Université d'Aix-Marseille*

Aurélie BANNAY
*Assistant Hospitalo-Universitaire en épidémiologie, économie de la santé et prévention*
*UFR de Médecine, Université de Lorraine, Nancy*

Sophie BASTIDE
*Assistant Hospitalo-Universitaire en épidémiologie, économie de la santé et prévention*
*UFR de Médecine, Université Montpellier-Nîmes*

Lisa BELIN
*Biostatisticienne*
*Faculté de médecine Pierre et Marie Curie, Paris*

Paul de BOISSIEU
*Assistant Hospitalo-Universitaire en épidémiologie, économie de la santé et prévention*
*UFR de Médecine, Université Paris Sud*

Aurélie BOURMAUD
*Praticien Hospitalier en Santé Publique*
*Centre Hygée, Institut de Cancérologie Lucien Neuwirth, Saint Priest en Jarez*

Bastien BOUSSAT
*Assistant Hospitalo-Universitaire en épidémiologie, économie de la santé et prévention*
*UFR de Médecine, Université Grenoble Alpes*

Laurent BOYER
*Professeur en épidémiologie, économie de la santé et prévention*
*UFR de médecine, Université d'Aix-Marseille*

Aurélie DAUMAS
*Maître de Conférences en thérapeutique*
*UFR de médecine, Université d'Aix-Marseille*



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Agnès DECHARTRES
*Maître de Conférences en épidémiologie, économie de la santé et prévention*
*Faculté de médecine Paris Descartes*

Anne-Marie DOLS
*Assistant Hospitalo-Universitaire en épidémiologie, économie de la santé et prévention*
*UFR de Médecine, Université Grenoble Alpes*

Antoine DUCLOS
*Professeur en épidémiologie, économie de la santé et prévention*
*Faculté de Médecine Lyon Est*

Pascale FABBRO-PERAY
*Maître de Conférences en épidémiologie, économie de la santé et prévention*
*UFR de Médecine, Université Montpellier-Nîmes*

Sarah FELDMAN
*Interne*
*Faculté de médecine Pierre et Marie Curie, Paris*

Adeline GALLINI
*Maître de Conférences en épidémiologie, économie de la santé et prévention*
*UFR Toulouse Rangueil, Université de Toulouse*

Virginie GARDETTE
*Maître de Conférences en épidémiologie, économie de la santé et prévention*
*UFR Toulouse Rangueil, Université de Toulouse*

Adrien GUILLOTEAU
*Assistant hospitalo-Universitaire en épidémiologie, économie de la santé et prévention*
*UFR des Sciences de la Santé, Université de Bourgogne Franche-Comté, Dijon*

David HAJAGE
*Praticien Hospitalier en Santé Publique*
*Hôpital La Pitié Salpêtrière, Assistance Publique – Hôpitaux de Paris*

Marie HERR
*Maître de Conférences en épidémiologie, économie de la santé et prévention*
*UFR des Sciences de la Santé Simone Veil, Université Versailles Saint-Quentin-en-Yvelines*

Jérémie JEGU
*Maître de Conférences en épidémiologie, économie de la santé et prévention*
*UFR Médecine, Université de Strasbourg*

Damien JOLLY
*Professeur en épidémiologie, économie de la santé et prévention*
*UFR de médecine, Université de Reims Champagne Ardenne*

Lukshe KANAGARATNAM
*Assistant Hospitalo-Universitaire en épidémiologie, économie de la santé et prévention*
*UFR de médecine, Université de Reims Champagne Ardenne*



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

José LABARERE
*Professeur en épidémiologie, économie de la santé et prévention*
*UFR de Médecine, Université Grenoble Alpes*

Guy LAUNOY
*Professeur en épidémiologie, économie de la santé et prévention*
*UFR Santé, Université de Caen »*

Benoît MARIN
*Professeur en épidémiologie, économie de la santé et prévention*
*UFR de Médecine, Université de Limoges*

Rémy de MIL
*Assistant Hospitalo-Universitaire en épidémiologie, économie de la santé et prévention*
*UFR Santé, Université de Caen »*

Pierre NAZEYROLLAS
*Professeur en Thérapeutique*
*UFR de médecine, Université de Reims Champagne Ardenne*

Pauline OCCELLI
*Praticien Hospitalier en Santé Publique*
*Hospices Civils de Lyon*

Mathieu ORIOL
*Assistant Hospitalo-Universitaire en épidémiologie, économie de la santé et prévention*
*UFR de Médecine, Université Jean Monnet de Saint-Etienne*

Nadia OUBAYA
*Assistant Hospitalo-Universitaire en épidémiologie, économie de la santé et prévention*
*UFR de Médecine, Université Paris-Est Créteil*

Pierre-Marie PREUX
*Professeur en épidémiologie, économie de la santé et prévention*
*UFR de Médecine, Université de Limoges*

Noémie RESSEGUIER
*Maître de Conférences en épidémiologie, économie de la santé et prévention*
*UFR de médecine, Université d'Aix-Marseille*

Arnaud SEIGNEURIN
*Maître de Conférences en épidémiologie, économie de la santé et prévention*
*UFR de Médecine, Université Grenoble Alpes*

Matthieu WARGNY
*Assistant Hospitalo-Universitaire en épidémiologie, économie de la santé et prévention*
*UFR de Médecine, Université de Nantes*



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Essais thérapeutiques/ Essais d'intervention

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 1 Essai de supériorité

Arnaud Seigneurin, Anne-Marie Dols

| | |
|---|---|
| **Titre de l'article** | **Randomized Trial of Radiation Therapy Plus Procarbazine, Lomustine, and Vincristine Chemotherapy for Supratentorial Adult Low-Grade Glioma : Initial Results of RTOG 9802** |
| **Auteurs** | Shaw EG, Wang M, Coons SW, Brachman DG, Buckner JC, Stelzer KJ, Barger GR, Brown PD, Gilbert MR, Mehta MP |
| **Référence** | J Clin Oncol. 2012 ; 30:3065-70. |
| **Lien vers l'article** | https://www. ncbi.nlm.nih.gov/pmc/articles/PMC3732006/ |

**1** **Il s'agit d'un essai**

a. De non-infériorité
b. Contre placebo
c. En groupes parallèles
d. En double insu
e. En cross-over

**2** **Par rapport à une randomisation simple, la méthode de randomisation utilisée dans l'essai permet de renforcer la limitation des déséquilibres entre le groupe de patients traités par radiothérapie seule et le groupe de patients traités par radiothérapie et chimiothérapie PCV concernant**

a. L'effectif de patients
b. La répartition par classe d'âge
c. La survie globale
d. La répartition par type histologique de gliome
e. La répartition par sexe

**3** **Le critère de jugement principal de l'essai**

a. Est un critère clinique
b. Est un critère composite
c. Est un critère de substitution
d. Est un critère intermédiaire
e. Correspond à un objectif thérapeutique



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **Le calcul du nombre de sujets nécessaire dans l'essai est basé sur**

a. La valeur de la survie sans progression à 5 ans dans le groupe traité par radiothérapie
b. Une augmentation absolue de 15 % de la survie globale à 5 ans dans le groupe traité par radiothérapie et chimiothérapie PCV comparativement à celui traité par radiothérapie seule
c. Une augmentation relative de 21 % de la survie globale à 5 ans dans le groupe traité par radiothérapie et chimiothérapie PCV comparativement à celui traité par radiothérapie seule
d. Un seuil alpha global de 5 %
e. Une valeur élevée de puissance statistique

**5** **L'analyse statistique du critère de jugement principal**

a. A été conduite en intention de traiter
b. Est à l'origine d'un biais de classement
c. Vise à tester une hypothèse de supériorité
d. Favorise la démonstration de l'hypothèse testée
e. Présente une inflation du risque alpha

**6** **Les données des figures 2 et 3 du suivi des 2 groupes au cours de la période de 8 ans permettent**

a. D'affirmer que le nombre de perdus de vus est inférieur à 5 %
b. D'affirmer que le nombre de progressions tumorales est inférieur au nombre de décès
c. D'affirmer que le nombre de décès est supérieur dans le groupe traité par radiothérapie seule
d. De calculer la médiane de survie sans progression dans le groupe traité par radiothérapie associée à la chimiothérapie PCV
e. D'affirmer que 17 décès par gliome de bas grade sont survenus dans le groupe traité par radiothérapie seule au cours des 2 premières années

**7** **La survie globale à 5 ans ne diffère pas de façon statistiquement significative entre le groupe traité par radiothérapie seule et le groupe traité par radiothérapie associée à la chimiothérapie PCV car**

a. Les courbes de survie globale dans chacun des 2 groupes obtenues par la méthode de Kaplan-Meier se croisent
b. La médiane de survie globale n'a pas été atteinte pour le groupe traité par radiothérapie et chimiothérapie PCV
c. L'intervalle de confiance à 95 % du hazard ratio comprend la valeur 1
d. La valeur p du test du log-rank est de 0,13
e. Les taux de mortalité instantanés entre les 2 groupes au cours des 5 années de suivi sont comparables



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**8** **L'absence de différence statistiquement significative entre les 2 groupes de l'essai pour la survie globale à 5 ans peut être expliquée par**

a. Un nombre de décès trop faible pour atteindre une puissance statistique de 90 %
b. Un manque de puissance statistique pour pouvoir mettre en évidence une augmentation de la survie globale à 5 ans de 9 points dans le groupe traité par radiothérapie et chimiothérapie PCV
c. L'absence d'efficacité de la radiothérapie
d. Un déséquilibre de répartition des classes d'âge < 40 ans et ≥ 40 ans dans les 2 groupes
e. L'absence d'efficacité de la chimiothérapie par PCV

**9** **L'analyse statistique conduite chez les patients qui ont survécu 2 ans indique que**

a. La survie globale à 7 ans est supérieure pour les patients traités par radiothérapie et chimiothérapie PCV
b. La survie globale à 7 ans est de 74 % dans le groupe traité par radiothérapie et chimiothérapie PCV
c. Le type histologique de la tumeur est un facteur de confusion dans l'association entre la survie à 7 ans et le groupe de traitement
d. La survie sans progression à 7 ans est supérieure dans le groupe traité par radiothérapie et chimiothérapie PCV
e. Les progressions tumorales après 2 années de suivi ont été peu fréquentes dans le groupe traité par radiothérapie seule

**10** **Le croisement des courbes de survie globale des deux groupes de l'essai (figure 2) peut être expliqué par**

a. L'effet de la chimiothérapie PCV sur la survie globale qui n'est pas constant au cours du temps
b. La présence de la radiothérapie comme facteur de confusion entre la chimiothérapie PCV et le critère de jugement principal
c. L'effet de la chimiothérapie PCV sur la survie globale qui est retardé de plusieurs années
d. Une interaction entre la durée de suivi et l'effet de la chimiothérapie PCV sur la survie globale
e. Les réponses A, B, C, D sont fausses

**11** **La conclusion des auteurs indiquant que la survie sans progression est améliorée chez les patients traités par radiothérapie et chimiothérapie PCV par rapport au traitement par radiothérapie seule est à prendre avec précaution car elle est basée sur une analyse**

a. Présentant une augmentation du risque alpha
b. En sous-groupe
c. D'un critère de jugement secondaire
d. Présentant une puissance statistique faible
e. Per protocole



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**12** **L'analyse du critère de jugement concernant la survie sans progression présente une inflation du risque alpha en raison de**

a. La réalisation d'analyses intermédiaires
b. La réalisation d'analyses ajustées
c. L'utilisation d'un critère de jugement intermédiaire
d. La répétition de tests statistiques
e. La présence d'un biais de classement dans la mesure de la progression tumorale

**13** **Les points forts de la méthodologie de cet essai incluent**

a. Une randomisation décentralisée
b. Une analyse en intention de traiter
c. Une puissance statistique élevée
d. Un critère de jugement principal clinique
e. Un groupe contrôle pertinent

**14** **Les limites méthodologiques de cet essai pour l'estimation de l'effet de la chimiothérapie PCV sur la survie globale incluent**

a. La possibilité de réaliser des traitements de sauvetage en cas de progression tumorale
b. Le caractère ouvert de l'essai
c. L'utilisation d'un critère de jugement de substitution comme critère de jugement principal
d. L'absence de prise en compte de l'extension de la résection tumorale comme facteur de confusion
e. Le nombre de patients inclus dans l'essai

**15** **Vous pouvez affirmer à partir des résultats de cet essai que**

a. Les traitements par radiothérapie seule et par radiothérapie associée à la chimiothérapie PCV ont une efficacité équivalente sur la survie globale à 5 ans des gliomes de bas grade réséqués
b. La survie sans progression à 5 ans des gliomes de bas grade réséqués est améliorée par la chimiothérapie PCV associée à la radiothérapie
c. Tous les patients atteints d'un gliome de bas grade devraient être traités par résection chirurgicale, radiothérapie et chimiothérapie PCV
d. La chimiothérapie par temozolomide doit être associée à la résection chirurgicale et à la radiothérapie en remplacement de la chimiothérapie PCV dans le traitement des gliomes de bas grade
e. Les réponses A, B, C, D sont fausses



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 1

**Réponse exacte : C**

*Commentaires* Il s'agit d'un essai randomisé de supériorité en 2 groupes parallèles visant à démontrer la supériorité de l'association radiothérapie et chimiothérapie par procarbazine, lomustine et vincristine (PCV) sur la radiothérapie seule dans le traitement des gliomes de bas grade après résection chirurgicale. L'essai n'est pas conduit en double aveugle puisque les participants et les médecins ont connaissance du bras d'affectation. Un essai en cross-over consiste à prendre chaque participant comme son propre témoin : un groupe reçoit successivement le traitement A puis le traitement B, tandis que l'autre groupe reçoit le traitement B puis le traitement A.

## Question 2

**Réponses exactes : A – B – D**

*Commentaires* Il s'agit d'une randomisation stratifiée et permutée par bloc. La randomisation par bloc permet de limiter les déséquilibres d'effectif entre les 2 groupes en tirant au sort des blocs de patients comprenant chacun une répartition égale des 2 types de traitement comparés. La stratification permet de contrôler la distribution aléatoire des types de traitements dans chacune des strates (par exemple < 40 ans et ≥ 40 ans), de façon à équilibrer les 2 groupes au sein de chaque strate. Dans l'essai la randomisation était stratifiée par classe d'âge, par score de performance Karnofsky, par la présence d'un rehaussement à l'IRM pré opératoire et par type histologique. Par contre la survie globale correspond au critère de jugement principal de l'essai.

## Question 3

**Réponses exactes : A – E**

*Commentaires* Le critère de jugement principal de cet essai est un critère clinique correspondant à la survie globale (overall survival), c'est-à-dire à la survie quelle que soit la cause de décès. Il s'agit d'un critère qui correspond directement à l'objectif thérapeutique d'augmenter la survie des patients.

Un critère composite prend en considération simultanément plusieurs types d'événements. Le critère de jugement secondaire est la survie sans progression qui est un critère composite puisqu'il prend en compte deux évènements : le décès quelle que soit la cause ou la progression de la tumeur.

Un critère de substitution est un examen paraclinique ou un signe physique utilisé à la place d'un critère clinique, sachant qu'une variation du critère de substitution induite par le traitement étudié est supposée refléter la variation du critère clinique. Un critère de jugement intermédiaire ne correspond pas directement à un objectif thérapeutique mais documente plutôt les mécanismes d'action du traitement.

## Question 4

**Réponses exactes : B – C – D – E**

*Commentaires* Le nombre de sujets nécessaire a été calculé en se basant sur une puissance statistique de 90 %, ce qui correspond à une valeur élevée (la puissance correspond à la probabilité de mettre en évidence une efficacité supérieure de la radiothérapie associée à la chimiothérapie PCV si cette association est réellement plus efficace). Le calcul du nombre de sujets nécessaire se base également sur une survie globale à 5 ans de 70 % dans le groupe traité par radiothérapie seule et de 85 % dans le groupe traité par radiothérapie et chimiothérapie PCV, ce qui correspond à une augmentation absolue de 15 % de la survie (85 %-70 %) et à une augmentation relative de 21 % (85 %/70 % = 1,21), ainsi que sur un risque alpha global de 5 %.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 5

**Réponses exactes : A – C**

*Commentaires* La figure 1 indique que 251 patients ont été randomisés, 126 dans le groupe traité par radiothérapie seule et 125 dans le groupe traité par radiothérapie et chimiothérapie PCV. Toujours selon cette figure, la totalité de ces patients ont été analysés dans leur groupe de randomisation alors que la compliance au traitement variait de 95 % à 97 % selon la définition retenue, ce qui correspond à une analyse en intention de traiter. L'analyse statistique ne peut pas entrainer de biais de classement qui est défini par la présence d'une erreur dans la mesure du critère de jugement de l'essai. L'analyse est conduite pour tester une différence de survie entre 2 groupes, ce qui correspond à une hypothèse de supériorité. Dans le cadre des essais de supériorité, l'analyse en intention de traiter est plus conservatrice par rapport à l'analyse per protocole. En effet, l'analyse per protocole ne conserve que les patients qui se sont conformés au protocole, ce qui peut favoriser la démonstration d'une supériorité. L'analyse du critère de jugement principal est réalisée sans inflation du risque alpha puisqu'il est précisé que les 3 analyses intermédiaires ont été réalisées en s'assurant que le risque alpha global était de 5 %.

## Question 6

**Réponse exacte : C**

*Commentaires* La figure 2 indique que 51 et 37 décès sont survenus respectivement dans les groupes traités par radiothérapie seule, et par radiothérapie associée à la chimiothérapie PCV. Le nombre de personnes à risque pour chaque année de suivi est indiqué sur cette figure et on constate qu'il y avait bien 126 et 125 personnes à risque en début de suivi, ce qui correspond au nombre de patients inclus, mais seules 2 et 4 personnes étaient à risque à 8 ans. Le nombre de décès n'explique pas la différence entre l'effectif de départ et les 6 personnes à risque à 8 ans et la différence correspond aux perdus de vue qui sont supérieurs à 5 %.

Sur la figure 3 le nombre d'événements rapportés (76 et 48) correspond aux décès + progressions (ne comptent que pour 1 événement les patients qui ont progression puis décès), on ne peut donc pas en déduire le nombre total de progressions. La seule chose que l'on peut déduire de ces 2 figures, c'est qu'il y a 51 décès et 25 progressions sans décès (76-51) dans le groupe RT seule et 37 décès et 11 progressions sans décès (48-37) dans le groupe RT + PCV, ce qui ne permet pas de répondre à la question. Mais combien de progressions parmi les décès : les figures ne permettent pas de le savoir. Il est donc impossible d'après les figures de comparer le nombre de décès et le nombre total de progression.

La médiane de survie sans progression correspond à la durée à laquelle 50 % des patients ont subi l'événement considéré (décès toute cause ou progression tumorale). Son calcul n'est pas possible pour le groupe traité par radiothérapie associée à la chimiothérapie PCV car la survie sans progression à 8 ans est supérieure à 50 %, ce qui implique que moins de la moitié des patients ont subi l'événement considéré à la fin du suivi.

Les décès rapportés lors du suivi ne concernent que les décès toutes causes, et le nombre de décès par gliome de bas grade n'est pas rapporté dans l'article.

## Question 7

**Réponses exactes : C – D – E**

*Commentaires* Les courbes de survie obtenues dans chacun des 2 groupes se croisent et cela signifie que l'effet de la chimiothérapie PCV n'est pas constant au cours du suivi. Par contre ce croisement ne permet pas de conclure à l'absence de différence statistiquement significative



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

de survie dans les 2 groupes. La médiane de survie globale n'a pas été atteinte pour le groupe traité par radiothérapie et chimiothérapie PCV mais cela n'a pas de rapport avec la comparaison des 2 groupes.

Le test du log-rank permet de tester l'hypothèse d'une différence des courbes de survie obtenues par la méthode de Kaplan-Meier. Ce test a une valeur p supérieure au seuil choisi ce qui permet de conclure à l'absence de différence statistiquement significative de survie globale entre les 2 groupes. Le fait que l'intervalle de confiance à 95 % du hazard ratio comprenne la valeur 1 implique que les taux de mortalité instantanés ne sont pas différents entre les 2 groupes, et donc, qu'une différence de survie globale entre les deux groupes n'a pas pu être mise en évidence. Il faut noter que le croisement des courbes de survie des 2 groupes de cet essai invite à la prudence quant à l'interprétation des résultats des tests de comparaison. Ces derniers doivent en effet être réalisés en l'absence de croisement des courbes. C'est pourquoi les auteurs ont réalisé une seconde analyse parmi les patients qui ont survécu au moins 2 ans pour se placer dans une situation ou les courbes de survie ne se croisent pas.

## Question 8

**Réponses exactes : B – E**

*Commentaires* La puissance statistique est liée au nombre de décès dans chaque groupe qui dépend du nombre de patients dans l'essai et de leur survie. L'essai a inclus 251 patients ce qui est conforme au calcul du nombre de sujets nécessaire (120 patients par groupe) pour mettre en évidence une augmentation relative de 21 % de la survie globale à 5 ans (de 70 % à 85 %) avec une puissance statistique de 90 % et un risque alpha global de 5 %. Les résultats de l'essai montrent une survie globale à 5 ans de 63 % et 72 % respectivement pour les groupes radiothérapie seule, et radiothérapie plus chimiothérapie par PCV. La survie à 5 ans est donc plus basse dans l'essai par rapport aux hypothèses du calcul du nombre de sujets nécessaire, ce qui correspond à un nombre plus important de décès. Un trop faible nombre de décès ne peut donc pas expliquer un manque de puissance statistique. Par contre, le nombre de sujets de l'essai a été calculé pour mettre en évidence une augmentation de 15 points de la survie globale à 5 ans (de 70 % à 85 %) et l'essai présente une puissance statistique plus faible pour mettre en évidence une augmentation de la survie inférieure à 15 points. La radiothérapie n'est pas le traitement étudié puisqu'elle fait partie du traitement de chacun des groupes. Son potentiel manque d'efficacité ne peut donc pas expliquer l'absence de différence statistiquement significative entre les 2 groupes de l'essai pour la survie globale à 5 ans. La randomisation a été stratifiée sur les classes d'âge $< 40$ ans et $\geq 40$ ans ce qui garantit un équilibre dans la répartition de ces 2 classes d'âge dans les 2 groupes de traitement. Par contre, l'absence d'efficacité de la chimiothérapie par PCV est une explication potentielle à l'absence de différence statistiquement significative entre les 2 groupes de l'essai pour la survie globale à 5 ans.

## Question 9

**Réponses exactes : A – B – D**

*Commentaires* Chez les patients qui ont survécu 2 ans, la survie globale à 7 ans est de 74 % dans le groupe traité par radiothérapie et chimiothérapie PCV et de 59 % dans le groupe traité par radiothérapie seule. Avec un hazard ratio de 0,52 ($IC_{95\,\%}$ : 0,30 – 0,90) et une p valeur du test du log-rank de 0,02, la survie globale à 7 ans est supérieure pour les patients traités par radiothérapie et chimiothérapie PCV par rapport à ceux traités par radiothérapie seule. L'analyse



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

ajustée sur le type histologique retrouve un hazard ratio identique de 0,52 ($IC_{95\ \%}$ : 0,30 – 0,91) par rapport à l'analyse non ajustée, ce qui signifie que le type histologique n'est pas un facteur de confusion dans la relation entre la survie et le groupe de traitement.

Si le texte précise que 65 progressions tumorales ont été constatées parmi les 126 patients du groupe traité par radiothérapie seule (paragraphe « Treatment of tumor recurrence and failure patterns »), le nombre de progressions tumorales parmi le groupe des patients ayant survécu au moins 2 années n'est en revanche pas précisé, ni la date de survenue de ces progressions tumorales. Il n'est donc pas possible d'affirmer que les progressions tumorales après 2 années de suivi ont été peu fréquentes dans le groupe traité par radiothérapie seule.

### Question 10

**Réponses exactes : A – C – D**

*Commentaires* Le croisement des 2 courbes de survie environ 2,5 années après l'inclusion dans l'essai témoigne d'une interaction entre l'effet de la chimiothérapie PCV et le temps, c'est-à-dire que l'effet n'est pas constant au cours du temps et il pourrait être retardé de plusieurs années. La radiothérapie n'est pas un facteur de confusion qui pourrait expliquer ce croisement des courbes de survie. Au cours des 2 premières années, la courbe de survie du groupe traité par radiothérapie et chimiothérapie PCV se situe en dessous de la courbe de survie du groupe traité par radiothérapie seule, ce qui implique que la chimiothérapie PCV n'augmente pas la survie globale dans les 2 premières années.

### Question 11

**Réponses exactes : A – C**

*Commentaires* Les auteurs ont contrôlé l'inflation du risqua alpha pour l'analyse du critère de jugement principal. L'analyse de la survie sans progression correspond à l'analyse d'un critère de jugement secondaire pour lequel il n'y a pas de contrôle du risque alpha. Il ne s'agit pas d'une analyse per protocole mais d'une analyse en intention de traiter car la figure 1 indique que l'ensemble des patients ont été analysés dans leur groupe de randomisation, alors que la compliance au traitement variait de 95 % à 97 % selon la définition retenue. Il ne s'agit pas d'une analyse de sous-groupe puisque l'analyse concerne l'ensemble des participants à l'essai. La question d'une trop faible puissance statistique se pose lorsqu'on n'observe pas de différence statistiquement significative entre les 2 groupes, ce qui n'est pas le cas pour le résultat sur la survie sans progression.

### Question 12

**Réponse exacte : D**

*Commentaires* L'analyse du critère de jugement concernant la survie sans progression présente une inflation du risque alpha car il existe une répétition d'un test statistique au seuil de 5 % sans procédure de contrôle du risque alpha. Un tel contrôle aurait pu être réalisé, par exemple, en abaissant le seuil de signification statistique retenu pour les analyses. L'analyse de la survie sans progression n'est pas une analyse intermédiaire dans l'essai et la survie sans progression n'est pas un critère de jugement intermédiaire. La réalisation d'analyses ajustées vise à prendre en compte l'existence de facteurs de confusion mais n'est pas l'explication d'une augmentation du risque alpha. Enfin, un biais de classement pourra avoir comme conséquence un biais dans l'estimation de l'effet de la chimiothérapie PCV mais n'entraine pas d'augmentation du risque alpha.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 13

**Réponses exactes : B – C – D – E**

*Commentaires* La décentralisation de la randomisation n'est pas un point fort d'un essai puisqu'on lui préfère une randomisation centralisée qui permet de garantir l'imprévisibilité de la randomisation. Par ailleurs, l'article ne précise pas si une randomisation centralisée a été réalisée. L'analyse de l'essai a été conduite en intention de traiter ce qui constitue un point fort dans les essais de supériorité puisqu'elle préserve la comparabilité initiale des 2 groupes issue de la randomisation. Les investigateurs ont basé le calcul du nombre de sujets nécessaire sur une puissance statistique de 90 % ce qui est une valeur élevée. Le critère de jugement principal est la survie globale qui est un critère clinique. C'est un point fort de l'essai car cela correspond directement à un objectif thérapeutique contrairement aux critères intermédiaires qui documentent plutôt les mécanismes d'actions du traitement. Le groupe contrôle est constitué des patients traités par radiothérapie seule ce qui est un groupe contrôle pertinent puisqu'il correspond au traitement habituel après la résection chirurgicale.

## Question 14

**Réponses exactes : A – B**

*Commentaires* Un traitement de sauvetage est autorisé lors de la survenue d'une progression tumorale pour des raisons éthiques évidentes, ce qui pourrait minimiser l'effet de la chimiothérapie PCV sur la survie globale. En effet, davantage de patients ont présenté des progressions tumorales dans le groupe radiothérapie seule (52 %) par rapport au groupe radiothérapie et chimiothérapie PCV (30 %), ce qui a abouti à davantage de traitements de sauvetage dans le groupe radiothérapie seule. L'analyse qui est réalisée en intention de traiter va prendre en compte l'effet des traitements de sauvetage dans la comparaison entre les 2 groupes, ce qui pourrait minimiser l'effet du traitement par chimiothérapie PCV. Par contre, l'analyse de la survie sans progression s'affranchit de ce problème en étudiant la survenue des décès et des progressions tumorales. Le caractère ouvert de l'essai est une limite par rapport à un essai en double aveugle qui permet de minimiser les différences de suivi et d'évaluation dans les 2 groupes. Le critère de jugement principal est la survie globale incluant les décès quelle que soit leur cause, ce qui ne correspond pas à un critère de substitution. Pour rappel, un critère de substitution est un critère intermédiaire dont le changement induit par le traitement reflète un changement du critère clinique final. L'intérêt d'utiliser un critère de substitution dans un essai est d'anticiper l'effet sur le critère clinique final avec une durée de suivi des sujets inférieure et/ou un nombre de sujets dans l'essai inférieur. L'extension de la résection tumorale est un facteur de confusion potentiel qui pourrait être associé à la survie globale à 5 ans. La prise en compte de ce facteur n'a pas été prévue lors de la stratification de la randomisation. En revanche un modèle de Cox ajusté sur plusieurs variables dont l'extension de la résection tumorale a été réalisé, ce qui a permis de prendre en compte ce facteur de confusion. Le nombre de patients inclus est conforme au calcul du nombre de sujets nécessaire afin d'obtenir une puissance statistique de 90 % pour mettre en évidence une augmentation relative de la survie globale de 21 %, ce qui est une valeur pertinente.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 15

**Réponse exacte : E**

*Commentaires* L'absence de différence statistiquement significative en termes de survie globale dans les 2 groupes obtenue par cet essai de supériorité ne permet pas d'affirmer l'équivalence d'efficacité des 2 stratégies thérapeutiques. C'est un essai d'équivalence basé sur une méthodologie particulière qui permettrait de répondre à la question d'une efficacité équivalente. L'amélioration statistiquement significative de la survie sans progression à 5 ans du groupe traité par radiothérapie et chimiothérapie PCV est constatée sur l'analyse d'un critère de jugement secondaire avec un risque d'inflation du risque alpha. L'essai ne montre pas de différence concernant le critère de jugement principal retenu (différence de survie globale à 5 ans lors de l'ajout d'une chimiothérapie PCV) : on ne peut donc pas affirmer avec les résultats de cet essai qu'il faut traiter les patients atteints de gliome de bas grade par résection chirurgicale, radiothérapie et chimiothérapie PCV. L'essai ne permet pas de comparer la chimiothérapie PCV et la temozolomide. L'efficacité de la temozolomide associée à la radiothérapie est une question de recherche qui nécessiterait de nouveaux essais.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 2 Essai de supériorité

Aurélie Daumas

| | |
|---|---|
| **Titre de l'article** | **Transfusion strategies for acute upper gastrointestinal bleeding** |
| **Auteurs** | Villanueva C, Colomo A, Bosch A, Concepción M, Hernandez-Gea V, Aracil C, Graupera I, Poca M, Alvarez-Urturi C, Gordillo J, Guarner-Argente C, Santaló M, Muñiz E, Guarner C |
| **Référence** | N Engl J Med. 2013 ; 368(1):11-21. |
| **Lien vers l'article** | https://www.ncbi.nlm.nih.gov/pubmed/23281973 |

**1** **L'objectif principal de cette étude est**

a. D'évaluer si l'incidence cumulée des complications transfusionnelles au cours des hémorragies digestives diminue en fonction d'une stratégie de transfusion libérale ou restrictive
b. D'évaluer si la prévalence des complications transfusionnelles au cours des hémorragies digestives diminue en fonction d'une stratégie de transfusion libérale ou restrictive
c. D'évaluer si la stratégie restrictive est associée à moins de décès que la stratégie libérale au cours des hémorragies digestives
d. D'évaluer si un seuil restrictif de transfusion de culots globulaires chez les patients qui présentent une hémorragie digestive est plus efficace qu'un seuil de transfusion libéral
e. D'estimer la fréquence des accidents transfusionnels lors de la transfusion des hémorragies digestives selon que la transfusion ait été décidée en fonction d'un seuil libéral ou restrictif.

**2** **Il s'agit d'une étude**

a. Observationnelle
b. En double aveugle
c. Randomisée
d. Multicentrique
e. De supériorité

**3** **La randomisation de la stratégie transfusionnelle**

a. Vise à réduire la durée de suivi des sujets
b. Sert à prendre en compte les fluctuations d'échantillonnage
c. Doit avoir lieu juste avant l'intervention proposée
d. A pour but de constituer des groupes comparables à l'inclusion
e. Garantit la comparabilité des groupes au moment de l'analyse



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **Concernant l'allocation de la stratégie transfusionnelle**

a. Elle permet d'améliorer la comparabilité initiale des groupes sur l'existence ou non d'une cirrhose
b. Elle permet de garantir le double aveugle
c. Elle est réalisée après vérification de l'éligibilité des patients
d. Elle permet de garantir un suivi identique des 2 groupes
e. Elle utilise des enveloppes opaques, scellées et numérotées

**5** **Quel(s) est (sont) le(les) point(s) fort(s) de la randomisation dans cette étude ?**

a. La stratification sur l'existence d'une cirrhose
b. La stratification sur le centre
c. Une randomisation par une procédure informatisée
d. Une randomisation par blocs
e. Une randomisation à l'aide d'enveloppes scellées

**6** **Le calcul du nombre de sujets nécessaires est basé sur**

a. Une puissance statistique de 95 %
b. Un risque de première espèce de 10 %
c. Un risque de 2e espèce à 10 %
d. Une différence de mortalité de 5 points en pourcentage entre les 2 groupes
e. Une borne de la zone d'équivalence de 5 %

**7** **Les auteurs ont choisi pour critère de jugement principal ?**

a. Un critère unique
b. Un critère subjectif
c. Un critère binaire
d. Un critère composite
e. Un critère biologique

**8** **Dans cette étude**

a. 921 patients ont été évalués pour éligibilité dans l'étude
b. 648 patients présentant une hémorragie digestive n'ont pu être inclus du fait des critères de non-inclusion de l'étude
c. 962 patients pouvaient participer à l'étude
d. L'analyse a été faite en intention de traiter modifiée
e. 67 patients cirrhotiques étaient child-pugh A

**9** **Concernant les résultats de l'étude**

a. 61 patients ont été transfusés dans le groupe stratégie de transfusion libérale
b. 47 patients ont été transfusés pour un autre motif que le taux d'hémoglobine
c. Il n'y a pas de différence significative concernant la quantité de cristalloïdes perfusée entre les 2 groupes
d. Il y a autant de patients qui ont reçu moins de 3 que plus de 3 culots globulaires dans le groupe stratégie de transfusion libérale
e. Il y a significativement eu moins de recours à un tips (transjugular intrahepatic porto-systemic shunt) dans le groupe stratégie de transfusion restrictive



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**10** **Au terme de l'étude, la survie à 45 jours est**

a. De 9 % dans le groupe stratégie de transfusion libérale
b. Significativement différente entre les 2 groupes
c. De 95 % dans le groupe stratégie de transfusion restrictive
d. Significativement différente entre les 2 groupes chez les patients cirrhotiques
e. Significativement meilleure en cas d'ulcère peptique dans le groupe stratégie de transfusion restrictive

**11** **À propos de la figure 2, parmi les propositions suivantes, laquelle ou lesquelles est (sont) exacte(s) ?**

a. Les traits horizontaux représentent les intervalles de confiance à 95 % dans la figure 2B
b. A j15, 404 patients initialement inclus dans le groupe stratégie transfusionnelle restrictive sont décédés
c. Dans le sous-groupe cirrhose, la stratégie transfusionnelle restrictive ne diminue pas significativement la mortalité à j45 (S6)
d. Le faible effectif peut expliquer que la mortalité ne diffère pas significativement entre les 2 stratégies transfusionnelles pour le sous-groupe saignement sur varices
e. Dans le sous-groupe saignement sur varices, le risque instantané de décès est diminué de 42 % avec la stratégie transfusionnelle par rapport à la stratégie libérale

**12** **Quelle(s) est (sont) la (les) limite(s) de l'étude ?**

a. L'exclusion des patients les plus graves
b. Le choix d'une analyse en intention de traiter
c. L'exclusion de certains patients de l'analyse
d. Le caractère monocentrique
e. Le choix du critère principal

**13** **La survenue plus fréquente des complications dans le groupe « transfusion libre » peut être expliquée par**

a. Une surcharge volémique
b. Des mécanismes immuno-allergiques
c. Le nombre de culots globulaires transfusés
d. Le nombre d'ulcères gastriques responsables du saignement
e. Le nombre de patients cirrhotiques

**14** **En conclusion, les résultats de cette étude**

a. Sont difficilement applicables à la population cible en raison de l'absence de précision sur les caractéristiques démographiques et les comorbidités
b. Sont applicables à tous les patients cirrhotiques
c. Doivent être vérifiés par une étude en double aveugle
d. Peuvent entrainer un changement de pratique
e. Sont cliniquement pertinents



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**15** **Pour appliquer les résultats de cette étude à vos patients, la(les) donnée(s) manquante(s) est (sont) ?**

a. Le pourcentage de patients coronariens
b. Le pourcentage de patients avec une insuffisance cardiaque
c. L'âge moyen des patients
d. L'origine géographique des patients
e. La gravité des patients



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 1**

**Réponse exacte : C**

**Inadmissible : D**

*Commentaires* La réponse se trouve dans le paragraphe Outcomes measures and definitions : « The primary outcome measure was the rate of death from any cause within the first 45 days... ».

**Question 2**

**Réponses exactes : C – E**

**Inadmissibles : A – B – D**

*Commentaires* Il s'agit d'un essai thérapeutique contrôlé, conduit en 2 groupes parallèles de façon prospective, randomisé, en ouvert, de supériorité à l'hôpital Santa Creu i San Pau à Barcelone.

**Question 3**

**Réponses exactes : C – D**

**Indispensable : D**

**Inadmissibles : A – B – E**

*Commentaires* Pour éviter un biais de sélection différentiel, la nature du traitement que reçoit un patient ne doit dépendre d'aucun facteur susceptible d'influencer le résultat. Ainsi, l'allocation du traitement ne doit dépendre ni de la gravité de la maladie, ni des caractéristiques du patient, ni du contexte des soins donnés au patient. Seule une attribution au hasard peut garantir la totale indépendance de la nature du traitement donné vis-à-vis de ces facteurs. Autrement, si un traitement est donné plus spécifiquement à des patients d'un certain type, les groupes ne seront plus comparables et il y aura biais. Pour obtenir cette indépendance les traitements sont attribués aléatoirement aux patients. Ainsi, chaque patient, quelles que soient ses caractéristiques, a la même probabilité de recevoir un traitement ou l'autre. La randomisation crée deux groupes de patients comparables en moyenne.

Les fluctuations aléatoires d'échantillonnage peuvent faire apparaître entre les groupes une différence qui en réalité n'existe pas. Dans une situation où le risque est identique dans les 2 groupes, par hasard, le pourcentage observé dans un groupe pourra être inférieur à ce qu'il aurait dû être tandis que dans l'autre groupe, le hasard conduit à une valeur observée surestimant la vraie valeur. À l'opposé, les fluctuations aléatoires peuvent aussi faire disparaître une différence qui existe pourtant. C'est la réalisation d'un test statistique qui permet de conclure, à partir de l'observation, sur l'existence (ou non) d'une vraie différence entre les deux groupes. Devant une différence observée, le test statistique permet de calculer la probabilité que l'on aurait d'observer ce résultat si en réalité il n'y avait pas de différence entre les deux groupes. Cette probabilité est appelée p. Classiquement, il a été convenu que le risque acceptable d'erreur alpha est de 5 %. Ainsi, devant une différence observée, on conclura à l'existence d'une réelle différence seulement si le risque de se tromper pris en faisant cette conclusion est inférieur à 5 %, c'est-à-dire, si la valeur de p donnée par le test est inférieure au seuil de 5 %.

**Question 4**

**Réponses exactes : A – C – E**

**Indispensable : C**

**Inadmissible : B**

*Commentaires* La randomisation a été stratifiée sur la présence ou non d'une cirrhose permettant d'améliorer la comparabilité initiale des groupes sur cette variable et de supprimer l'effet du facteur de confusion qu'est la présence ou non d'une cirrhose.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

La randomisation a reposé sur l'utilisation d'enveloppes scellées ayant pour inconvénient notamment de permettre (par fraude) une prise de connaissance anticipée de la nature du traitement que vont recevoir les futurs patients. Il devient alors possible d'arranger le moment ou un patient sera inclus afin qu'il reçoive le traitement souhaité. Seule une procédure informatisée de randomisation garantit suffisamment son imprévisibilité. Le numéro ou la nature de traitement est obtenu par appel d'un centre de randomisation et après enregistrement du patient en train d'être inclus dans l'essai.

La randomisation doit avoir lieu après vérification de l'éligibilité du patient et juste avant le début de la prise en charge pour éviter au maximum que des sujets randomisés finalement ne participent pas à l'étude par retrait de consentement ou survenue d'un évènement les empêchant de poursuivre l'étude par exemple...

Le double aveugle évite toutes différences dans le suivi et l'évaluation des deux groupes. Le principe du double insu consiste à faire que tous les patients, quelle que soit leur appartenance à l'un des groupes de l'essai, apparaissent identiques. Ceci est obtenu en ne révélant pas la nature exacte du traitement reçu par les patients. La randomisation ne garantit ni le double aveugle ni un suivi identique des 2 groupes, la preuve il s'agit ici d'une étude en ouvert malgré une randomisation de la stratégie transfusionnelle.

### Question 5

**Réponses exactes : A – D**

**Inadmissible : B – C**

*Commentaires* *Cf.* Question 4 pour la stratification sur la cirrhose et la randomisation par enveloppes scellées.

Il s'agit ici d'un essai monocentrique sans stratification sur le centre mais une stratification sur le centre fréquemment réalisée dans les essais multicentriques consiste à équilibrer le nombre de patients recevant la stratégie ou le traitement A et recevant B au sein de chaque centre. Elle permet de minimiser l'effet centre qui découle du fait que le recrutement et la prise en charge varient selon le centre.

L'intérêt de randomiser par blocs est de limiter au maximum les déséquilibres entre les groupes, puisqu'approximativement autant de patients vont recevoir chaque traitement. Cependant pour garantir une imprévisibilité totale, l'utilisation de blocs de taille différente est recommandée.

### Question 6

**Réponse exacte : D**

**Inadmissibles : A – B – E**

*Commentaires* Il s'agit d'une étude de supériorité.

Les paramètres à prendre en compte pour le calcul du nombre de sujets nécessaires avec un critère de jugement principal qui n'est pas une variable continue sont :

- le risque alpha ;
- le risque bêta ;
- la différence attendue entre les 2 groupes ;
- et la valeur attendue de la variable dans le groupe témoin et éventuellement le pourcentage de patients inanalysables.

Dans l'étude la puissance statistique est de 80 % car bêta = 1- puissance = 1-0,2 = 0,8 ou 80 %, le risque de première espèce est de 5 % et une différence de mortalité de 5 % entre les 2 groupes est suspectée en prenant une mortalité à 10 % dans le groupe stratégie transfusionnelle libérale.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 7**

**Réponses exactes : A – C**

**Inadmissibles : B – E**

*Commentaires* Le critère de jugement est le taux de décès toutes causes dans les 45 premiers jours suivant la transfusion. Il est unique. Il n'est pas composite puisqu'il n'a pas plusieurs composantes. Il s'agit bien évidement d'un critère clinique, objectif et pertinent. Le critère de jugement correspond ici à une variable qualitative binaire ou dichotomique (survenue ou non de l'évènement = décès ou survie).

**Question 8**

**Réponses exactes : B – C – D – E**

**Indispensable : E**

**Inadmissible : A**

*Commentaires* *Cf.* Flow-chart et tableau 1

1610 Patients ont été screenés c'est-à-dire présentaient les critères d'inclusion. 962 étaient incluables, c'est-à-dire ne présentaient pas un ou plusieurs des critères de non-inclusion donc sous réserve de leur consentement pouvaient être inclus dans l'étude. Sur les 962 incluables seulement 921 ont accepté de participer à l'étude et ont été randomisés.

L'analyse en ITT doit tenir compte de tous les patients randomisés or dans l'étude, seulement 889 patients sur les 921 randomisés ont été analysés. Selon le tableau 1, 67 patients étaient child-pugh A (37 et 30 respectivement dans chaque groupe), 155 étaient child-pugh B (76 et 79) et 55 étaient child-pugh C (26 et 29).

**Question 9**

**Réponses exactes : B – C – D – E**

*Commentaires* 61 patients n'ont pas été transfusés dans le groupe transfusion libre *versus* 225 dans le groupe transfusion restrictive (item A faux). Il n'y a effectivement pas de différence significative entre les 2 groupes sur la quantité de cristalloïdes perfusés (item C faux).

Concernant le nombre de CGR transfusés, la médiane est de 3 dans le groupe transfusion libérale. La médiane permet de couper l'ensemble des valeurs en deux parties égales : mettant d'un côté une moitié des valeurs, qui sont toutes inférieures ou égales à 3 et de l'autre côté l'autre moitié des valeurs, qui sont toutes supérieures ou égales à 3.

**Question 10**

**Réponses exactes : B – C**

**Inadmissible : A**

*Commentaires* Attention, la question concerne la survie or le critère de jugement est le taux de décès. Cependant, comme il y a une différence significative sur le taux de décès entre les 2 groupes, il y a une différence évidemment sur la survie qui est significativement meilleure dans le groupe transfusion libérale. Il n'y a pas de différence significative entre les 2 stratégies transfusionnelles si on compare tous les patients cirrhotiques. Par contre, si on compare uniquement les patients child-pugh A et B, on note une survie significativement meilleur pour les patients du groupe transfusion restrictive.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 11**

**Réponses exactes : A – C – D – E**

**Indispensable : A**

**Inadmissible : B**

*Commentaires* Dans la figure 2B, l'Hazard ratio est représenté sous la forme d'un rectangle vertical. La barre horizontale qui l'encadre est l'IC à 95 %. Si cette droite horizontale comprend l'ordonnée 1 cela revient à dire que 1 est dans l'intervalle de confiance et donc que l'on n'a pas montré que l'Hazard ratio était différent de 1 au seuil de risque alpha de 5 %. Les résultats dans les sous-groupes de gravité de la cirrhose peuvent être non significatifs du fait d'un manque de puissance au vu du faible effectif dans ces sous-groupes.

**Question 12**

**Réponses exactes : A – C – D**

**Indispensables : A – C**

*Commentaires* L'analyse qui doit être utilisée pour les études de supériorité est l'analyse en intention de traiter. Tous les patients inclus sont analysés dans leur groupe de randomisation quoi qu'il arrive. À l'opposé, l'analyse per protocole est le gold standard pour les études de non-infériorité. Dans cette analyse, seuls sont étudiés les patients ayant eu le bon protocole de traitement du début à la fin sans écart au protocole.

Le caractère monocentrique et le fait en plus que le centre soit en Espagne peut faire doute de l'applicabilité des résultats en France. De plus, du fait des critères de non inclusion nombreux et restrictifs, les patients les plus graves ou fragiles ne pouvaient pas participer à l'étude limitant l'applicabilité des résultats observés aux patients vus en pratique courante, très souvent plus graves.

**Question 13**

**Réponses exactes : A – B – C**

**Inadmissible : E**

*Commentaires* On note un nombre significativement différent de CGR transfusés entre les 2 groupes ; ceci peut expliquer, par des mécanismes de surcharge volémique ou immuno-allergiques, des effets indésirables plus fréquents dans le groupe « Transfusion libre ». Compte tenu du fait que le nombre d'ulcères responsables d'un saignement n'est pas significativement différent entre les 2 groupes, cela ne peut pas expliquer la différence observée en termes de complications. De même, il n'y a pas de différence significative entre les 2 groupes en termes de patients cirrhotiques.

**Question 14**

**Réponses exactes : A – D – E**

**Indispensable : A**

*Commentaires* Les résultats de cette étude ne sont pas applicables aux patients cirrhotiques tous confondus mais uniquement aux patients cirrhotiques respectant les critères d'inclusion et de non-inclusion de l'étude. La capacité à extrapoler les résultats d'un essai à la vraie vie dépend du degré de sélection des patients inclus dans l'étude.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 15**

**Réponses exactes : A – B – C**

**Indispensable : C**

*Commentaires* Le caractère monocentrique et étranger de l'étude limite l'applicabilité des résultats en France de façon générale. La connaissance de l'origine géographique des patients n'aidera pas à extrapoler les résultats à vos patients qui habitent en France. La gravité des patients est indiquée par le Rockall score. On n'a par contre aucune idée de l'âge des patients et de leurs comorbidités en dehors des comorbidités mentionnées dans les critères de non-inclusion. Ce manque d'information limite l'extrapolation des résultats.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 3 Essai de supériorité

Sophie Bastide, Pascale Fabbro-Peray

| | |
|---|---|
| **Titre de l'article** | **Epidural steroid injections compared with gabapentin for lumbosacral radicular pain : multicenter randomized double blind comparative efficacy study** |
| **Auteurs** | Cohen SP, Hanling S, Bicket MC, White RL, Veizi E, Kurihara C, Zhao Z, Hayek S, Guthmiller KB, Griffith SR, Gordin V, Anderson White M, Vorobeychik Y, Pasquina PF |
| **Référence** | BMJ 2015 ; 350:h1748 |
| **Lien vers l'article** | http://www.bmj.com/content/350/bmj.h1748  |

**1** **Quel était le contexte justifiant cette étude ?**

a. L'infiltration par corticostéroïdes est une stratégie thérapeutique qui a montré son efficacité pour le traitement de la douleur dans la lombosciatique
b. Le traitement médicamenteux par gabapentinoïdes est une stratégie thérapeutique qui a montré son efficacité pour le traitement de la douleur dans la lombosciatique
c. L'association d'un traitement médicamenteux par gabapentinoïdes et d'une infiltration par corticostéroïdes est une stratégie thérapeutique qui a montré son efficacité pour le traitement de la douleur dans la lombosciatique
d. Aucune étude randomisée n'a comparé l'infiltration par corticostéroïdes à un traitement médicamenteux dans la lombosciatique
e. Aucune étude randomisée n'a comparé l'infiltration par corticostéroïdes au traitement médicamenteux par gabapentinoïdes dans la lombosciatique

**2** **Les objectifs de cette étude étaient**

a. D'évaluer l'efficacité de l'association thérapeutique d'un traitement médicamenteux par gabapentine et d'un traitement par infiltration de corticostéroïdes sur la douleur de jambe et de dos dans la lombosciatique
b. De comparer l'efficacité d'un traitement médicamenteux par gabapentine *versus* un traitement par infiltration de corticostéroïdes sur la douleur de jambe et de dos dans la lombosciatique
c. D'évaluer les complications survenant avec un traitement médicamenteux par gabapentine par rapport à celles survenant avec un traitement par infiltration de corticostéroïdes dans la prise en charge de la lombosciatique
d. De comparer l'efficacité d'un traitement médicamenteux par gabapentine *versus* un traitement par infiltration de corticostéroïdes sur la qualité de vie dans la lombosciatique
e. De comparer l'efficacité d'un traitement médicamenteux par gabapentine *versus* un traitement par infiltration de corticostéroïdes sur le handicap fonctionnel dans la lombosciatique



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**3** **Il s'agissait d'une étude**

a. De non-infériorité
b. De cohorte
c. Contrôlée contre placebo
d. En double aveugle
e. Multicentrique

**4** **Le protocole d'étude standardisé prévoyait**

a. La possibilité d'infiltration bilatérale et/ou de plusieurs étages en fonction de la symptomatologie clinique
b. Un guidage échographique systématique de l'infiltration
c. Une confirmation IRM du diagnostic étiologique de la lombosciatique
d. Une prise en charge kinésithérapique standardisée concomitante de 6 semaines
e. La possibilité d'un traitement de secours médicamenteux en cas d'échec de la stratégie thérapeutique interventionnelle

**5** **Les critères de qualité de la randomisation dans cette étude étaient**

a. La stratification sur le centre
b. La stratification sur le type d'infiltration
c. Le ratio 1:1
d. La randomisation par blocs
e. La taille variable aléatoire des blocs

**6** **Les critères permettant d'assurer le double aveugle dans cette étude étaient**

a. L'encapsulation de la gabapentine et du placebo
b. L'infiltration par un médecin indépendant
c. L'utilisation des indices de James et de Bang (« James blinding index » et « Bang blinding index »)
d. La randomisation
e. L'utilisation d'un double placebo

**7** **Le critère de jugement principal était**

a. Un critère clinique
b. Un critère objectif
c. Un critère unique
d. Un critère de substitution
e. Un critère composite



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**8** **Concernant l'analyse statistique de cette étude**

a. Elle a été réalisée en intention de traiter pour l'ensemble des critères de jugement concernant la douleur
b. Différentes analyses en sous-groupes prévues au protocole ont été réalisées
c. Une correction de Bonferroni a été réalisée pour l'analyse principale du critère de jugement principal
d. Elle a été ajustée sur le sexe pour l'ensemble des critères de jugement concernant la douleur
e. Une méthode d'imputation des données manquantes par tirage au sort a été utilisée pour l'ODI (Oswestry disability index)

**9** **Le calcul du NSN (nombre de sujets nécessaire)**

a. Est présenté avec l'ensemble des informations nécessaires pour être vérifié
b. Se fondait sur le critère de jugement principal à 1 mois
c. Se basait sur une hypothèse d'efficacité plausible et justifiée
d. Prévoyait une marge de 13 % de sorties d'essai
e. Aurait donné un nombre plus important en supposant une différence de 2 points dans le score de douleur entre les deux groupes (IMMPACT guidelines)

**10** **D'après le tableau 1**

a. Il n'existait aucune différence notable entre les caractéristiques des patients des deux bras de l'étude
b. Les patients randomisés dans le bras de traitement par infiltration de corticostéroïdes, comparés à ceux randomisés dans le bras de traitement par gabapentine, étaient de manière plus importante des femmes (34 % *versus* 18 %)
c. Les patients randomisés dans le bras de traitement par infiltration de corticostéroïdes, comparés à ceux randomisés dans le bras de traitement par gabapentine, étaient beaucoup moins traités pas opioïdes à l'inclusion (74 % *versus* 7 %)
d. Les patients randomisés dans le bras de traitement par infiltration de corticostéroïdes, comparés à ceux randomisés dans le bras de traitement par gabapentine, présentaient une corpulence moyenne inférieure (IMC = 18 kg/m² vs 26 kg/m²)
e. Les différents scores de douleur étaient similaires dans les deux bras signifiant une absence de différence d'efficacité entre les deux traitements évalués

**11** **Les principaux résultats de l'étude sont**

a. Une absence de différence significative entre les deux bras pour l'ensemble des critères explorant la douleur de jambe
b. Une efficacité significativement plus importante des deux traitements sur la douleur de jambe que sur la douleur de dos
c. Un recours au traitement chirurgical de deuxième intention dans l'année comparable entre les deux bras
d. Une tolérance similaire entre les deux bras
e. Une meilleure efficacité globale perçue dans le bras avec traitement par infiltration de corticostéroïdes



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**12** **Dans cette étude**

a. La population source est représentative de la population cible
b. L'échantillon inclus est représentatif de la population cible
c. L'échantillon inclus est représentatif de la population source
d. L'échantillon analysé pour le critère de jugement principal est représentatif de l'échantillon inclus
e. Les résultats obtenus peuvent se généraliser à l'ensemble de la population cible

**13** **Dans cette étude peuvent être rencontrés**

a. Un biais d'attrition important qui limite l'interprétation de l'ensemble des critères de jugement
b. Un biais de classement différentiel sur les critères de jugement du fait d'un aveugle imparfait
c. Un biais de confusion sur le sexe lié à l'absence d'appariement
d. Un biais de sélection important lié au perdu de vue
e. Un biais de manque de puissance

**14** **Finalement, à partir de cette étude, il est possible de conclure que**

a. Le traitement par gabapentine est équivalent au traitement par infiltration de corticostéroïdes pour le traitement de la lombosciatique
b. Le traitement à préconiser en première intention dans la lombosciatique est l'infiltration par corticostéroïdes
c. Les deux traitements évalués ont une efficacité sur la douleur de la lombosciatique majeure et durable dans le temps
d. Les deux traitements évalués présentent une tolérance parfaite
e. De nouvelles études évaluant l'efficacité de l'association de traitements sont nécessaires

**15** **Globalement**

a. La validité interne de l'étude est assurée
b. La validité externe de l'étude est assurée
c. Il ne semble pas avoir de conflits d'intérêts majeurs limitant la validité des résultats
d. La grille de lecture adaptée pour cet article est la grille CONSORT
e. L'étude a suivi les obligations réglementaires et de bonne pratique clinique



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 1

**Réponses exactes : A – B – E**

*Commentaires*

A. Vrai : l'efficacité contre placebo de l'infiltration par corticostéroïdes sur la douleur dans la lombosciatique avait déjà été prouvée (page 1, 2e colonne, dernier paragraphe/page 2, 1re colonne, 1er et 2e paragraphes). À noter que ces études concluaient à une ampleur d'effet modeste.

B. Vrai : l'efficacité contre placebo du traitement médicamenteux oral par gabapentinoïdes sur la douleur dans la lombosciatique avait également déjà été prouvée (page 2, 1re colonne, 1er et 2e paragraphes). À noter que ces études concluaient à une ampleur d'effet modeste.

C. Faux : l'efficacité contre placebo de l'association d'un traitement médicamenteux oral par gabapentinoïdes et d'une infiltration par corticostéroïdes sur la douleur dans la lombosciatique ne semble pas avoir été montrée.

D. Faux : différentes études avaient déjà comparé l'infiltration par corticostéroïdes à un traitement médicamenteux oral (tramadol, myorelaxants, AINS) (page 2, 1re colonne, 3e paragraphe). À noter qu'il s'agissait d'études de faible effectif et réalisées en ouvert.

E. Vrai : la comparaison thérapeutique de l'infiltration par corticostéroïdes au traitement médicamenteux par gabapentinoïdes avait déjà été testée dans la radiculopathie cervicale mais non dans la lombosciatique (page 2, 1re colonne, 3e paragraphe).

## Question 2

**Réponses exactes : B – C – E**

*Commentaires* L'objectif principal de cette étude était de comparer l'efficacité thérapeutique, sur la douleur moyenne de jambe dans la lombosciatique, d'un traitement médicamenteux par gabapentine *versus* un traitement par infiltration unique de corticostéroïdes.

Les objectifs secondaires étaient de comparer ces deux stratégies thérapeutiques en termes de :

- douleur maximale de jambe ;
- douleur moyenne de dos ;
- douleur maximale de dos ;
- handicap fonctionnel (questionnaire ODI : Oswestry disability index) ;
- complications et effets secondaires ;
- consommation d'antalgiques ;
- efficacité globale perçue.

## Question 3

**Réponses exactes : D – E**

**Inadmissible : B**

*Commentaires* Il s'agissait d'un essai thérapeutique (étude expérimentale), de différence d'efficacité, multicentrique internationale (7 centres : 3 centres militaires américains, 1 centre militaire européen, 3 centres civils américains), contrôlé, randomisé, en double aveugle, à deux bras parallèles (traitement médicamenteux par gabapentine *versus* traitement par infiltration unique de corticostéroïdes). Le bras « contrôle » n'était pas constitué d'un placebo mais chacun des deux bras comparés nécessitait l'utilisation d'un placebo pour permettre le double aveugle.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 4

**Réponses exactes : C – E**

*Commentaires*

A. Faux : l'infiltration pouvait bien être bilatérale par la voie interlaminaire mais était limitée dans le cadre de l'étude à un seul étage (page 3, 1re colonne, 1er paragraphe). Dans tous les cas, il ne pouvait avoir qu'une seule injection mais celle-ci pouvait diffuser.

B. Faux : il s'agissait d'un guide radiographique (page 2, 2e colonne, 3e paragraphe).

C. Vrai (page 2, 1re colonne, 5e paragraphe).

D. Faux.

E. Vrai : un traitement de secours médicamenteux par tramadol, AINS ou opioïdes pouvait être utilisé en cas d'échec de la stratégie thérapeutique interventionnelle (page 3, 1re colonne, 4e paragraphe).

### Question 5

**Réponses exactes : A – B – D**

*Commentaires* (Description de la randomisation page 2, 2e colonne, 2e paragraphe).

A. Vrai : la stratification sur les facteurs majeurs connus pour influencer l'efficacité d'un ou des traitements considérés est un critère de qualité de la randomisation. Dans une étude multicentrique, la stratification sur le centre est obligatoire pour prendre en compte l'éventuel « effet centre ».

B. Vrai : le type d'infiltration est un facteur prépondérant dans l'efficacité de l'infiltration, en fonction de l'uni ou de la bilatéralité de la douleur ; la stratification sur cette variable était donc importante.

C. Faux : la randomisation dans cette étude a bien été réalisée avec un ratio de 1:1, c'est une caractéristique de la randomisation à décrire, mais cela ne constitue pas un critère de bonne ou de mauvaise qualité de celle-ci.

D. Vrai : la randomisation par blocs permet de garantir un équilibre entre les bras lors d'une randomisation stratifiée (déséquilibre au maximum de la moitié de la taille du plus grand bloc).

E. Faux : dans cette étude, la taille des blocs est fixe par centre, mais juste variable entre les différents centres. Une randomisation par blocs avec des blocs de taille variable suivant, en plus, une séquence tirée au sort aléatoirement sont les caractéristiques assurant la meilleure qualité à la randomisation quel que soit le type d'essai clinique.

Remarque : la centralisation et l'indépendance de la personne en charge de la mise en œuvre sont d'autres critères de qualité d'une randomisation.

### Question 6

**Réponses exactes : A – B – E**

**Indispensable : E**

*Commentaires*

A. Vrai : l'encapsulation de la gabapentine et du placebo de traitement médicamenteux permettait de rendre les deux produits indiscernables, ce qui est obligatoire pour avoir un aveugle (page 3, 1re colonne, 3e paragraphe).

B. Vrai : la procédure d'infiltration étant légèrement différente en fonction du bras, l'infiltration par un médecin indépendant était obligatoire pour avoir un aveugle (page 3, 1re colonne, 2e paragraphe).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

C. Faux : les indices de James et de Bang permettent la vérification *a posteriori* de l'efficience de l'aveugle dans une étude mais ne sont pas des outils pour le garantir (page 3, 2e colonne, 4e paragraphe/page 4, 2e colonne, 1er paragraphe).

D. Faux : la randomisation et le double aveugle sont deux des trois points clefs méthodologiques des essais thérapeutiques de différence d'efficacité (avec l'analyse en intention de traiter) mais l'un ne garantit pas l'autre, et inversement. D'ailleurs, ils ne sont pas systématiquement associés dans les études. Une étude non randomisée peut bénéficier d'un suivi en aveugle et une étude en ouvert peut être randomisée.

E. Vrai : l'étude visant à comparer un traitement par voie injectable et un traitement per os, l'utilisation d'un double placebo était obligatoire pour garantir l'aveugle. Chaque patient inclus dans l'étude devait bénéficier de l'administration de deux produits, un actif et un placebo selon le bras (infiltration active et traitement placebo *versus* infiltration placebo et traitement actif).

## Question 7

**Réponses exactes : A – D**

*Commentaires* Le critère de jugement principal était la douleur moyenne de jambe sur la dernière semaine, mesurée à l'aide d'une échelle numérique simple entre 0 et 10, à 1 mois et à 3 mois (page 3, 1re colonne, 4e paragraphe).

A. Vrai.

B. Faux : la douleur exprimée est un critère subjectif.

C. Faux : la douleur est relevée à deux horizons temporels (1 mois et 3 mois).

D. Vrai : les auteurs définissent le critère de jugement final comme la douleur neuropathique mais elle n'est pas directement évaluable. La douleur de jambe est un critère de substitution pour celle-ci (page 3, 2e colonne, 2e paragraphe).

E. Faux. Un critère de jugement composite est un critère de jugement constitué par l'association de plusieurs critères d'importance clinique équivalente dont la présence d'un seul suffit pour considérer le critère composite comme rempli.

## Question 8

**Réponses exactes : A – B**

*Commentaires*

A. Vrai : tous les sujets inclus dans l'étude ont été analysés dans le groupe dans lequel ils ont été randomisés pour l'ensemble des critères de jugement concernant la douleur. En cas de données manquantes, les données ont été imputées pour inclure tous les sujets dans l'analyse (page 3, 2e colonne, 4e paragraphe/tableau 2, page 5).

B. Vrai : des analyses en sous-groupes ont été réalisées sur différentes variables démographiques et cliniques (étiologie de la lombosciatique, niveau de l'infiltration, ancienneté de la douleur, type d'infiltration, statut tabagique, statut militaire, présence de comorbidité psychiatrique, obésité, âge, sexe, dose de gabapentine) (page 4, 2e colonne, 1er paragraphe).

C. Faux : la correction de Bonferroni a été utilisée pour les analyses en sous-groupes (multiples analyses), pas pour l'analyse du critère de jugement principal qui était une comparaison unique (page 4, 2e colonne, fin du 1er paragraphe).

D. Faux : l'analyse des critères de jugement concernant la douleur a été ajustée sur le niveau de douleur de base grâce à l'analyse par ANCOVA, pas sur le sexe (page 3, 2e colonne, 5e paragraphe/tableau 2, page 5). Ce sont les analyses post hoc qui ont été ajustées sur le sexe.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

E. Faux : les données manquantes sur l'ODI (Oswestry disability index) ont été imputées en utilisant la dernière mesure disponible (page 3, 2e colonne, 4e paragraphe/tableau 2, page 5).

## Question 9

**Réponses exactes : B – C – D**

*Commentaires* Pour le calcul dans une étude du nombre de sujets nécessaire (NSN) pour la comparaison de deux moyennes observées, les informations nécessaires sont (*cf.* Description du NSN page 4, 2e colonne, 2e paragraphe) :

- le risque alpha = risque de 1re espèce ;
- la puissance souhaitée (1 – risque bêta = 1 – risque de 2e espèce) ;
- la différence de moyenne attendue entre les deux groupes ;
- l'écart-type attendu dans les deux groupes ;
- le caractère uni ou bilatéral de l'hypothèse alternative du test ;
- le ratio entre les deux groupes.

A. Faux : il manque l'information sur la puissance souhaitée pour l'étude (usuellement de 80 % à 90 %).

B. Vrai (pour l'horizon temporel à 1 mois).

C. Vrai : 2 références bibliographiques (42 et 43) sont utilisées pour justifier la différence choisie.

D. Vrai : « a retention rate of 87 % ». Les sorties prématurées d'essai incluent les éventuels perdus de vue, décès, retraits.

E. Faux : faire l'hypothèse d'une différence plus importante entre les deux groupes, toute autre information égale par ailleurs, donne un résultat de NSN plus faible.

## Question 10

**Réponse exacte : B**

**Indispensable : B**

*Commentaires*

A. Faux : il existait au moins une différence notable entre les patients des deux bras de l'étude concernant le sexe.

B. Vrai.

C. Faux : il s'agit d'une « coquille » manifeste dans le tableau. La proportion de patients sans traitement opioïde à l'inclusion était similaire dans les deux groupes, respectivement 74 % dans le bras de traitement par infiltration de corticostéroïdes et 76 % (et non 7 %) dans le bras de traitement par gabapentine.

D. Faux : la caractéristique de corpulence de la population incluse est exprimée dans l'étude avec une variable binaire « obésité » présentée en effectif et pourcentage (non en variable quantitative).

E. Faux : le tableau 1 présente les différents scores de douleur à l'inclusion, non les scores pour les critères de jugement, soit les scores après traitement. Les scores de base étaient similaires signant la bonne qualité de la randomisation pour la répartition des patients inclus en fonction de la gravité de leur atteinte.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 11

**Réponses exactes : C – D**

**Inadmissible : B**

*Commentaires*

A. Faux : une différence statistiquement significative a été mise en évidence pour la douleur maximale de jambe à 1 mois en faveur du traitement par infiltration de corticostéroïdes, différence égale à 0,9 avec un IC à 95 % de (0,0 à 1,9) et une p-value de 0,04 (tableau 2, page 5).

B. Faux : que ce soit en termes de douleur moyenne ou en termes de douleur maximale, à 1 mois ou à 3 mois, la diminution des scores de douleur semble effectivement plus importante pour la douleur de jambe que pour la douleur de dos, et cela pour les deux bras (tableau 2, page 5). Mais aucune analyse statistique formelle n'a été réalisée pour conclure sur ces comparaisons.

C. Vrai : aucune différence significative n'a été mise en évidence entre les deux bras de l'étude concernant le recours au traitement chirurgical de deuxième intention dans l'année, 13 % *versus* 14 % (p-value = 0,73) respectivement pour le traitement par infiltration unique de corticostéroïdes *versus* le traitement médicamenteux par gabapentine (tableau 3, page 6).

D. Vrai : aucune différence significative n'a été mise en évidence entre les deux bras de l'étude concernant les complications et effets secondaires (tableau 4, page 7/page 6, 1re colonne, 2e paragraphe). En termes d'injection, 8 % *versus* 10 % de complications (p-value = 0,75) ont été décrites, respectivement pour le traitement par infiltration unique de corticostéroïdes *versus* le traitement médicamenteux par gabapentine. En termes d'effets secondaires liés au traitement médicamenteux per os, 42 % *versus* 51 % d'effets secondaires (p-value = 0,24) ont été décrites, à nouveau respectivement pour le traitement par infiltration unique de corticostéroïdes *versus* le traitement médicamenteux par gabapentine.

E. Faux : aucune différence significative n'a été mise en évidence entre les deux bras de l'étude concernant l'efficacité globale perçue, que ce soit à 1 mois (67 % *versus* 57 % avec une p-value = 0,21) ou à 3 mois (45 % *versus* 33 % avec une p-value = 0,14) (toujours respectivement pour le traitement par infiltration unique de corticostéroïdes *versus* le traitement médicamenteux par gabapentine) (tableau 3, page 6).

## Question 12

**Réponse exacte : D**

**Indispensable : D**

*Commentaires*

A. Faux : la population cible est l'ensemble de la population adulte souffrant de lombosciatique. Les auteurs se placent dans le contexte « mondial » (page 1, 2e colonne, 1er paragraphe), mais leur justification correspond plus précisément aux pays industrialisés. La population source est beaucoup plus restreinte parce que le choix des centres se limite en grande majorité à des centres militaires, elle n'est donc pas représentative de la population cible.

B. Faux : si la population source n'est pas représentative de la population cible, l'échantillon d'étude, qui est directement issu dans celle-ci, ne peut pas être représentatif de la population cible.

C. Faux : il y a eu 34 % de refus de participation de patients qui remplissaient les critères d'éligibilité (74/(145 + 74)), c'est-à-dire de patients de la population source (figure, page 5). L'échantillon inclus n'est donc pas un échantillon aléatoire de la population source.

D. Vrai : il existe un biais d'attrition (biais de sélection) important ; les effectifs finaux à 3 mois étaient respectivement de 41/73 sujets dans le bras de traitement par infiltration de corticostéroïdes et de 32/72 dans le bras gabapentine (figure, page 5). Cependant, l'analyse statistique du



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

critère de jugement principal a été réalisée en intention de traiter avec imputation des données manquantes ; tous les sujets inclus dans l'étude ont été analysés dans le groupe dans lequel ils ont été randomisés. L'échantillon analysé pour le critère de jugement principal est identique à l'échantillon inclus, donc bien représentatif.

E. Faux : il existe un biais de sélection important lié au choix de la population source de l'étude. Les résultats obtenus ne peuvent pas se généraliser à l'ensemble de la population cible.

■ **Question 13**
**Réponse exacte : B**
**Indispensable : B**

*Commentaires*

A. Faux : les effectifs finaux à 3 mois étaient respectivement de 41/73 sujets dans le bras de traitement par infiltration de corticostéroïdes et de 32/72 dans le bras gabapentine, ce qui représente un biais d'attrition (biais de sélection) important (figure, page 5). Cependant, l'analyse statistique de la majorité des critères de jugement a été réalisée en intention de traiter avec imputation des données manquantes. L'interprétation de ces critères n'est donc pas biaisée par ce biais d'attrition (tableau 2, page 5).

B. Vrai : à 1 mois, une tendance à deviner le bon traitement a été mise en évidence avec le « Bang blinding index « dans le bras avec traitement médicamenteux par gabapentine. Si l'aveugle n'est pas maintenu, un risque de biais de classement existe surtout avec des critères de jugements subjectifs. Cette tendance n'étant retrouvé que dans un bras, le biais de classement risque de plus d'être différentiel.

C. Faux : le sexe peut certes être un facteur de confusion dans cette étude mais l'appariement n'est pas une méthode de prise en compte des facteurs de confusion possible car non adaptée aux essais thérapeutiques. (L'appariement est une méthode de prise en compte des facteurs de confusion utilisée dans les études observationnelles).

D. Faux : seulement 3 patients ont été perdus de vue dans cette étude durant le suivi (figure, page 5), le risque de biais de sélection dû à ceux-ci est donc négligeable.

E. Faux : un manque de puissance peut être discuté dans cette étude (*cf.* question 9) (page 7), ce qui est une limite de l'étude. Cependant un manque de puissance n'est pas un biais.

■ **Question 14**
**Réponse exacte : E**
**Indispensable : E**
**Inadmissible : A**

*Commentaires*

A. Faux : l'absence de différence significative ne signifie pas qu'il y ait une absence significative de différence.

B. Faux : aucune différence significative majeure n'a été mise en évidence entre les deux bras de l'étude permettant d'affirmer une telle conclusion (tableau 2, page 5/tableau 3, page 6/ tableau 4, page 7).

C. Faux (page 6, 2e colonne, 4e et 5e paragraphes).

D. Faux : 8 % à 10 % de complications ont été décrites pour l'infiltration et 42 % à 51 % d'effets secondaires pour le traitement médicamenteux per os (tableau 4, page 7/page 6, 1re colonne, 2e paragraphe).

E. Vrai (page 8, 1re colonne, 2e paragraphe).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 15

**Réponses exactes : C – D – E**

**Indispensable : D**

*Commentaires*

A. Faux (*cf.* question 13).

B. Faux (*cf.* question 12).

C. Vrai : les personnes impliquées dans l'étude (investigateur, personnel de recherche, statisticien) n'étaient pas directement en lien avec l'industrie pharmaceutique productrice des produits expérimentaux. Le financement de l'étude était institutionnel.

D. Vrai : la grille CONSORT est la grille de lecture adaptée aux essais cliniques.

E. Vrai : l'étude a été approuvée par un comité d'éthique indépendant, tous les sujets inclus ont signé un consentement de participation à l'étude (page 8, 2e colonne), et l'étude a bien été enregistrée sur ClinicalTrials.gov (page 1, fin du résumé). ClinicalTrials.gov est la base de données internationale d'enregistrement des essais cliniques permettant de rendre public de manière prospective le protocole d'une étude et ainsi de contribuer à la transparence de la réalisation de l'étude.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 4 Essai de supériorité

Pascale Fabbro-Peray, Sophie Bastide

| | |
|---|---|
| **Titre de l'article** | **A randomized, Controlled Trial of Total Knee Replacement** |
| **Auteurs** | Soren T.Skou, Ewa M ; Roos, Mogens B ; Laursen, Michael S. Rathleff, Lars Arendt-Nielsen, Ole Simonsen and Sten Rasmussen |
| **Référence** | N Engl J Med. 2015 ; 373:1597-606 |
| **Lien vers l'article** | http://www.nejm.org/doi/pdf/10.1056/NEJMoa1505467 |

**1** **Le(s) argument(s) des auteurs pour réaliser cette étude chez des sujets souffrant d'arthrose du genou est (sont)**

a. Un impact économique important dû à l'augmentation de la pratique chirurgicale de remplacement prothétique de genou
b. L'augmentation croissante des remplacements prothétiques de genou sans certitude que ce soit la meilleure prise en charge
c. Le manque de données fiables concernant l'intérêt de l'association du remplacement prothétique total de genou et d'une prise en charge non chirurgicale
d. Le manque de données fiables concernant l'efficacité des prises en charge non chirurgicales
e. Le manque de données fiables concernant la comparaison du remplacement prothétique total de genou aux prises en charge non chirurgicales

**2** **L'objectif de cette étude était de**

a. Montrer la supériorité sur l'amélioration des symptômes et de la qualité de vie d'un traitement chirurgical de remplacement prothétique du genou par rapport à un programme de prise en charge non chirurgical chez les patients souffrant d'arthrose du genou
b. Montrer la supériorité sur l'amélioration des symptômes et de la qualité de vie d'un traitement chirurgical de remplacement prothétique du genou suivi d'un programme de prise en charge non chirurgical par rapport aux soins usuels chez les patients souffrant d'arthrose du genou
c. Montrer la supériorité sur l'amélioration des symptômes et de la qualité de vie d'un traitement chirurgical de remplacement prothétique du genou suivi d'un programme de prise en charge non chirurgicale par rapport au programme de prise en charge non chirurgicale seule chez les patients souffrant d'arthrose du genou
d. Montrer l'intérêt sur l'amélioration des symptômes et de la qualité de vie d'un traitement chirurgical de remplacement prothétique du genou chez des patients souffrant d'arthrose du genou ayant suivi préalablement un programme de prise en charge non chirurgical
e. Évaluer l'efficacité sur l'amélioration des symptômes et de la qualité de vie d'un traitement chirurgical de remplacement prothétique du genou en complément d'un programme de prise en charge non chirurgical chez les patients souffrant d'arthrose du genou



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**3** **Il s'agit d'une étude**

a. Interventionnelle
b. Expérimentale
c. En aveugle de l'évaluateur
d. D'épidémiologie clinique
e. D'évaluation d'une intervention à multiples composantes

**4** **Dans cette étude, le programme de prise en charge non chirurgicale**

a. Était réalisé de façon standardisée chez les patients des deux groupes
b. Était personnalisé en fonction de certaines caractéristiques des patients
c. Comportait systématiquement des conseils diététiques
d. Comportait systématiquement des séances d'éducation thérapeutique
e. Comportait systématiquement l'administration de médicaments anti-douleurs

**5** **Quels sont les éléments de l'article permettant de caractériser la randomisation ?**

a. Elle est équilibrée
b. Elle a permis l'assignation secrète
c. Elle est stratifiée sur le centre
d. Elle est centralisée
e. Elle est réalisée de façon individuelle

**6** **Le critère de jugement principal est**

a. Clinique
b. Composite
c. Quantitatif continu
d. Objectif
e. Évalué en insu du groupe de traitement

**7** **Le fait qu'il n'y ait pas de placebo de l'intervention chirurgicale dans le groupe prise en charge non chirurgicale seule**

a. S'explique pour des raisons éthiques
b. Induit un risque de biais de classement différentiel sur le score de KOOS
c. Induit un risque de surestimation de l'effet de la chirurgie sur le score de KOOS
d. Peut induire une prise en charge non chirurgicale différente dans les deux groupes
e. Peut être à l'origine d'un biais de confusion

**8** **Si cette étude avait été réalisée en France en 2017, elle aurait nécessité**

a. L'avis consultatif d'un comité de protection des personnes
b. L'autorisation de l'Agence Nationale de Sécurité des Médicaments et des produits de santé (ANSM)
c. L'obtention du consentement oral des patients inclus
d. Le respect de la confidentialité des données des patients
e. La souscription d'une assurance par le promoteur de l'étude



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**9** **13 patients du groupe prise en charge non chirurgicale ont subi un remplacement prothétique du genou dans l'année suivant l'inclusion et 1 patient du groupe chirurgical n'a finalement pas été opéré. Cet écart au protocole**

a. Tend à sous-estimer l'effet de la chirurgie si on laisse les patients dans leur groupe de randomisation lors de l'analyse
b. Justifie l'exclusion de ces patients de l'analyse
c. Justifie que ces patients soient analysés selon le traitement reçu
d. Risque d'entraîner une perte de puissance statistique
e. Risque de biaiser le résultat dans l'analyse en intention de traiter sans qu'on ne puisse prévoir le sens du biais

**10** **D'après les résultats**

a. Il existe une différence moyenne significative de 16,5 points du score de KOOS à 12 mois entre le groupe traitement chirurgical et le groupe traitement non chirurgical
b. Il existe une amélioration moyenne significative de 16,5 points de la différence entre le score de KOOS à 12 mois et le score de KOOS à l'inclusion dans le groupe traitement chirurgical par rapport au groupe traitement non chirurgical
c. La différence sur le critère de jugement principal entre les deux groupes est significative car son intervalle de confiance à 95 % ne contient pas la valeur 1
d. Il y a 95 % de chances pour que la différence sur le critère de jugement principal entre les deux bras de traitement dans la population représentée par l'échantillon soit comprise entre 10,2 et 22,7
e. Dans les deux groupes, l'amélioration du score global de KOOS atteint la différence minimale cliniquement significative recommandée

**11** **Les résultats sur le critère de jugement principal d'efficacité**

a. Sont confortés par la concordance entre l'analyse brute et l'analyse ajustée
b. Sont confortés par la concordance entre l'analyse en intention de traiter et l'analyse per protocole
c. Ne sont interprétables que parce qu'il y a concordance entre l'analyse en intention de traiter et l'analyse per protocole
d. Ne sont interprétables que parce que la puissance de l'étude est élevée
e. Auraient nécessité que les auteurs prennent en compte l'inflation du risque alpha

**12** **Parmi les propositions suivantes concernant les résultats sur les critères de jugement secondaires, quelle(s) est (sont) celle(s) qui est (sont) exacte(s) ?**

a. Dans l'analyse en intention de traiter, les résultats sur les sous-scores de KOOS sont tous statistiquement significatifs
b. Dans l'analyse en intention de traiter, les résultats sur les sous-scores de KOOS pourraient s'expliquer par l'inflation du risque alpha
c. Dans l'analyse per protocole, les résultats sur les sous-scores de KOOS pourraient s'expliquer par l'inflation du risque alpha
d. Dans l'analyse per protocole, les résultats sur les sous-scores de KOOS pourraient s'expliquer par un manque de puissance
e. Dans l'analyse per protocole, les résultats pourraient s'expliquer par un biais de confusion



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**13** **Quels sont les moyens que les auteurs ont utilisés dans cette étude pour limiter les biais sur le recueil des critères de jugement ?**

a. L'aveugle de l'évaluateur
b. La standardisation de l'intervention non chirurgicale
c. L'allocation aléatoire des traitements
d. L'analyse en intention de traiter
e. La dissimulation par le patient de la zone chirurgicale du genou lors des évaluations

**14** **Suite à cette étude, pour les patients avec une arthrose confirmée modérée à sévère, éligible à un remplacement prothétique unilatéral du genou, sans antécédent de remplacement prothétique du même genou, on peut recommander**

a. Un remplacement prothétique du genou isolé
b. Un remplacement prothétique du genou suivi d'une prise en charge non chirurgicale multifacettes
c. Une prise en charge non chirurgicale multifacettes suivie d'un remplacement prothétique du genou
d. Une prise en charge non chirurgicale multifacettes seule si le patient le souhaite
e. Une ou plusieurs composantes de l'intervention non chirurgicale en fonction de la disponibilité du patient

**15** **Le(s) argument(s) limitant la généralisabilité des résultats est (sont)**

a. Le recrutement bicentrique des patients
b. Les recommandations de pratique sur lesquelles est basé le programme de prise en charge non chirurgicale
c. Le spectre de douleur des patients inclus
d. La proportion de patients éligibles incapables de participer au programme de prise en charge non chirurgicale
e. Le fait que les évaluations des critères de jugement aient été réalisées par un seul évaluateur



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 1

**Réponses exactes : A – B – E**

*Commentaires*

C. Faux : cet argument n'est pas cité par les auteurs.

D. Faux : des données d'efficacité des interventions non chirurgicales existent : Recent research has provided substantial evidence to suggest moderate effectiveness of nonsurgical treatments for knee osteoarthritis 6,7.

Les deux citations bibliographiques 6 et 7 font référence à des recommandations pour la prise en charge non chirurgicale de l'arthrose du genou.

## Question 2

**Réponses exactes : C – E**

*Commentaires*

D : Faux : la prise en charge non chirurgicale est réalisée après le remplacement prothétique du genou.

E : Vrai : c'est bien l'efficacité de la chirurgie dans une population soumise aussi à une prise en charge non chirurgicale.

## Question 3

**Réponses exactes : A – B – C – D**

*Commentaires*

A. Vrai : on évalue une intervention (la chirurgie prothétique).

Attention, le caractère « interventionnel » d'une étude n'a pas le même sens d'un point de vue épidémiologique et d'un point de vue réglementaire (dans la loi française). D'un point de vue épidémiologique, il s'agit d'étude où on va évaluer une « intervention » : qui peut être thérapeutique, de dépistage, de prévention, diagnostique, ou d'éducation pour la santé. D'un point de vue réglementaire, une étude va être considérée comme interventionnelle à partir du moment où les patients de l'étude vont être soumis à des procédures auxquelles ils ne seraient pas soumis s'ils n'étaient pas dans l'étude *et* qui leur font courir un certain « risque ». Exemple : dans une étude cas témoins à la recherche de facteurs de risque d'une maladie, le protocole d'étude impose la réalisation d'un prélèvement sanguin pour des dosages biologiques ; cette étude cas témoins, observationnelle d'un point de vue épidémiologique, sera classée en étude interventionnelle d'un point de vue réglementaire et imposera au promoteur de contracter une assurance. (*cf.* réponses à la question 8 pour classification réglementaire des études).

B. Vrai : tout essai thérapeutique RANDOMISE est expérimental. Les personnes qui reçoivent ou non l'intervention sont déterminés par l'évaluateur, aléatoirement.

Attention : Il existe une catégorie d'études entre les études par observation et les véritables études expérimentales : on les appelle les « études quasi-expérimentales ». Ces études comprennent une intervention, mais celle-ci n'est pas entièrement planifiée par la personne qui effectue la recherche.

Dans les études quasi-expérimentales, les personnes qui reçoivent ou non l'intervention sont déterminés par l'évaluateur de façon non aléatoire (pas de randomisation), mais elles doivent permettre d'identifier un groupe contrôle aussi proche que possible du groupe expérimental au niveau des caractéristiques initiales (préalables à l'intervention).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

C. Vrai : Ici seul l'évaluateurest en aveugle pour certains critères secondaires ; attention ce n'est pas un équivalent de simple aveugle qui est communément défini comme l'aveugle du patient.

D. Vrai : recherche chez les patients.

E. Faux : on évalue la chirurgie pas le programme de prise en charge non chirurgicale qui est réalisé chez les patients des deux groupes.

■ **Question 4**

**Réponses exactes : A – B – D**

*Commentaires*

A. Vrai : p1598, colonne 2 : To ensure proper standardization and to reduce the number of crossovers, the nonsurgical treatment was delivered to the two groups separately but identically, at the same facility, by specially trained physiotherapists and dietitians.

B. Vrai : pour certaines composantes de la prise en charge (diététique, gestion de la douleur).

C. Faux : seulement si BMI ≥ 25.

D. Vrai.

E. Faux : seulement si elle était nécessaire pour la participation à l'exercice physique.

■ **Question 5**

**Réponses exactes : A – E**

*Commentaires* Il y a très peu de détails sur la procédure de randomisation ; on ne peut affirmer qu'elle a permis l'assignation secrète des patients aux deux bras de traitement, ce qui est la qualité essentielle d'une bonne randomisation. On sait seulement que les patients sont randomisés individuellement selon un ratio 1:1 (c'est-à-dire de manière équilibrée entre les deux groupes).

Remarque : la randomisation peut se faire par cluster et non individuellement dans les essais en clusters (tous les sujets d'un même cluster reçoivent le même traitement).

■ **Question 6**

Réponses exactes : A – C

*Commentaires*

A. Vrai : objectif d'amélioration de symptômes et de qualité de vie.

B. Faux : un critère composite (« composite endpoint ») est un critère qui prend en considération simultanément plusieurs événementscliniques. Un patient présente le critère composite à partir du moment où il est victime de l'un des événements. La survenue successive de plusieurs des composantes du critère composite ne donne lieu qu'à une seule occurrence du critère composite.

C. Vrai : quantitatif (différence entre le score de KOOS à 12 mois et le score de KOOS de base), continu car le score de KOOS global pour un sujet est la moyenne de 4 sous scores.

D. Faux : subjectif ; il s'agit d'une autoévaluation par le patient.

E. Faux : autoévalué par le patient qui n'est pas en insu du groupe.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 7

**Réponses exactes : A – B – C – D – E**

*Commentaires* Tout est vrai. L'intervention placebo aurait permis l'aveugle du patient (mais n'est pas éthique) donc la diminution des biais de classement différentiel, la prise en compte de l'effet placebo lié à l'intervention, et la diminution du biais de confusion lié à l'absence d'aveugle (traitement non administré de la même façon, médicaments concomitants plus important dans un groupe que dans l'autre).

Le biais de classement est différentiel lorsque la mesure du critère de jugement peut engendrer une distorsion différente dans les deux groupes : ici les sujets qui ont subi la chirurgie peuvent évaluer les critères de façon plus favorable que ceux qui n'ont pas subi l'intervention (suggestion d'efficacité).

### Question 8

**Réponses exactes : B – D – E**

*Commentaires* Il s'agit d'une étude interventionnelle qui relève de la loi sur les recherches impliquant les personnes humaines (RIPH) entrée en vigueur en novembre 2016 (ex loi sur la recherche biomédicale) entrant dans la catégorie 1 (= recherche interventionnelle) qui nécessite donc l'accord d'un CPP (et non un avis consultatif), l'autorisation de l'ANSM, l'obtention du consentement écrit et non oral de chaque sujet inclus, le respect de la confidentialité du traitement des données et pour laquelle le promoteur doit toujours souscrire une assurance.

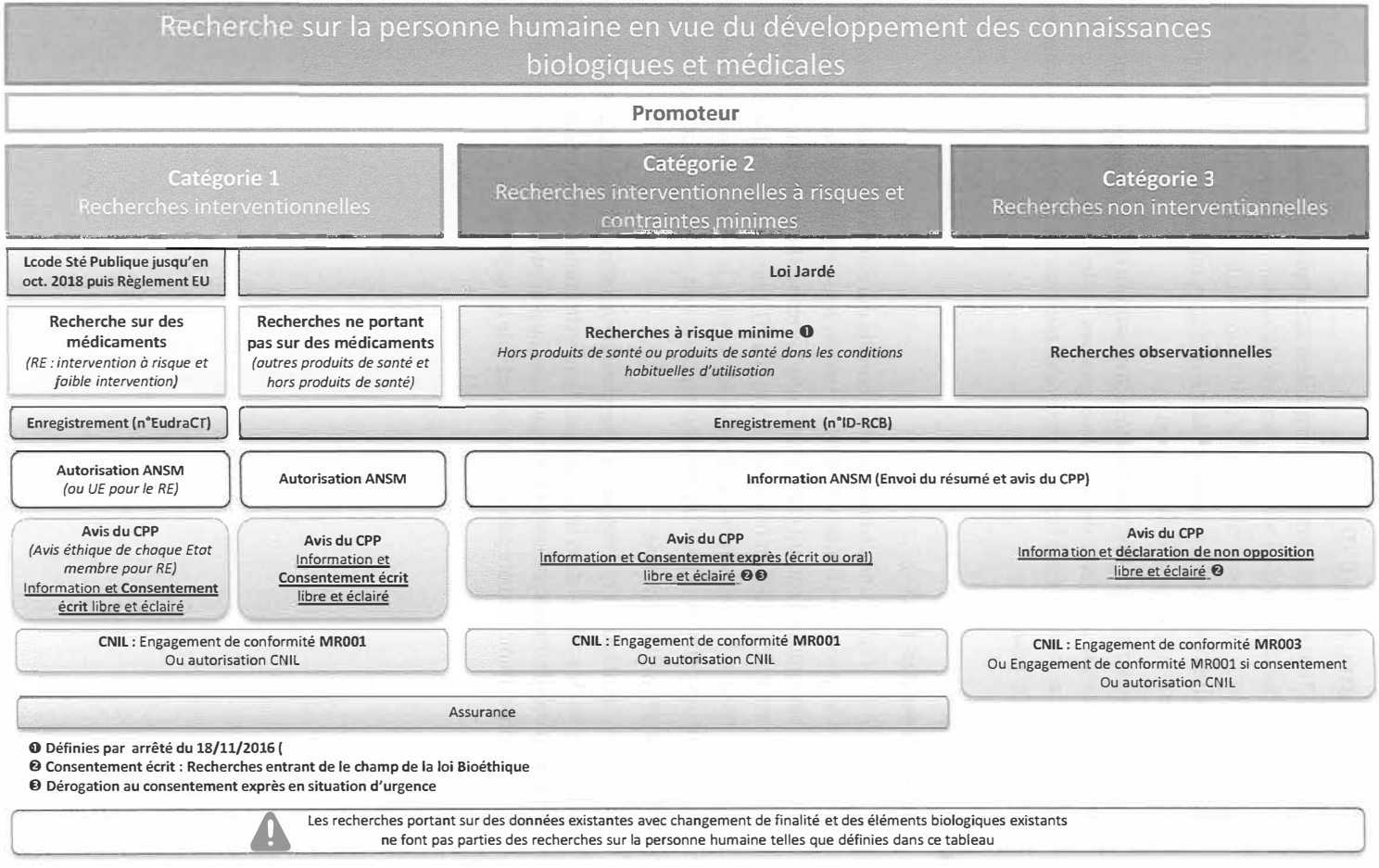
*Cf.* tableau récapitulatif ci-joint.

Remarque : la catégorie 2 représente la recherche interventionnelle avec risques et contraintes minimes pour le patient : dans la mesure où, dans cette étude, on évalue une chirurgie on ne peut pas parler de risques minimes. La liste des actes ou procédures considérés comme à risques ou contraintes minimes a été publiée dans le cadre de la loi JARDE.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# DÉMARCHES RÉGLEMENTAIRES EN FONCTION DU PROJET


Recherche sur la personne humaine en vue du développement des connaissances biologiques et médicales
Promoteur
Catégorie 1
Recherches interventionnelles
Catégorie 2
Recherches interventionnelles à risques et contraintes minimes
Catégorie 3
Recherches non interventionnelles
Lcode Sté Publique jusqu'en oct. 2018 puis Règlement EU
Loi Jardé
Recherche sur des médicaments
(RE : intervention à risque et faible intervention)
Recherches ne portant pas sur des médicaments
(autres produits de santé et hors produits de santé)
Recherches à risque minime ❶
Hors produits de santé ou produits de santé dans les conditions habituelles d'utilisation
Recherches observationnelles
Enregistrement (n°EudraCT)
Enregistrement (n°ID-RCB)
Autorisation ANSM
(ou UE pour le RE)
Autorisation ANSM
Information ANSM (Envoi du résumé et avis du CPP)
Avis du CPP
(Avis éthique de chaque Etat membre pour RE)
Information et Consentement écrit libre et éclairé
Avis du CPP
Information et Consentement écrit libre et éclairé
Avis du CPP
Information et Consentement exprès (écrit ou oral) libre et éclairé ❷❸
Avis du CPP
Information et déclaration de non opposition libre et éclairé ❷
CNIL : Engagement de conformité MR001
Ou autorisation CNIL
CNIL : Engagement de conformité MR001
Ou autorisation CNIL
CNIL : Engagement de conformité MR003
Ou Engagement de conformité MR001 si consentement
Ou autorisation CNIL
Assurance

❶ Définies par arrêté du 18/11/2016 (
❷ Consentement écrit : Recherches entrant de le champ de la loi Bioéthique
❸ Dérogation au consentement exprès en situation d'urgence

⚠ Les recherches portant sur des données existantes avec changement de finalité et des éléments biologiques existants ne font pas parties des recherches sur la personne humaine telles que définies dans ce tableau

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 9

**Réponses exactes : A – D**

*Commentaires*

A. Vrai : d'après l'hypothèse des auteurs, on s'attend avec le groupe chirurgie à une efficacité supérieure. Les patients finalement opérés dans le groupe non-chirurgie et non opérés dans le groupe chirurgie vont affaiblir la différence dans l'analyse donc sous-estimer l'effet de la chirurgie en intention de traiter.

B. Faux : analyse en intention de traiter indispensable pour limiter les biais.

C. Faux : *cf.* B.

D. Vrai : comme la différence va diminuer, la puissance statistique aussi.

E. Faux : *cf.* A.

### Question 10

**Réponses exactes : B – D – E**

*Commentaires*

A. Faux : il ne s'agit pas de la différence du score de KOOS à 12 mois mais de la différence de la différence de score entre l'inclusion et 12 mois.

B. Vrai (*cf.* commentaire A).

C. Faux : elle est significative car son intervalle de confiance à 95 % ne contient pas la valeur 0.

D. Vrai : définition de l'intervalle de confiance.

E. Vrai : l'augmentation du score global de KOOS (entre la mesure de base et la mesure au temps d'évaluation) cliniquement significative recommandée est de 10 points : page 1600 colonne 2. C'est vrai dans les deux groupes (16 points dans le groupe traitement non chirurgical et 32,5 dans le groupe traitement chirurgical) et les bornes inférieures des deux intervalles de confiance à 95 % sont supérieures à 10 (respectivement 10,1 dans le groupe traitement non chirurgical et 26,6 dans le groupe traitement chirurgical).

### Question 11

**Réponses exactes : A – B**

*Commentaires*

A. Vrai : dans un essai thérapeutique, l'analyse brute prévaut car il n'est en général pas nécessaire d'ajuster sur les facteurs de confusion ; ceux-ci étant pris en compte par la randomisation. Dans certains cas, un ajustement est réalisé soit pour prendre en compte des facteurs de confusion résiduels, soit pour améliorer la précision des estimations lorsqu'il y a eu stratification de la randomisation (on ajuste alors sur ces facteurs de stratification). Le fait que les deux analyses concordent conforte les résultats.

B. Vrai : dans un essai thérapeutique de différence, l'analyse en intention de traiter prévaut mais le fait que l'analyse en per protocole conduise à la même conclusion permet de penser qu'il n'y a pas eu trop d'écarts au protocole donc conforte les résultats.

C. Faux : dans un essai thérapeutique de différence, c'est toujours l'analyse en intention de traiter qui doit être prise en compte (elle permet de conserver la comparabilité des groupes car elle conserve dans l'analyse tous les patients randomisés dans leur groupe de randomisation y compris ceux qui ont eu des écarts au protocole, donc elle permet de se rapprocher de la « vraie vie » dans laquelle les patients sont moins compliants que dans un essai) ; l'analyse per



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

protocole, en excluant de l'analyse certains patients, est pourvoyeuse de biais. C'est seulement dans les études de non-infériorité ou d'équivalence que les deux analyses doivent concorder pour être interprétables (*cf.* cours sur les essais de non-infériorité).

D. Faux : le risque d'erreur intervenant pour interpréter un résultat significatif est le *risque alpha* (risque de conclure à une différence qui n'existe pas), pas le risque bêta (risque de ne pas conclure à une différence qui existe). La réflexion sur la puissance intervient seulement lorsque le résultat est non statistiquement significatif.

E. Faux : pas sur le critère de jugement principal (un seul test) ; en revanche cela aurait été nécessaire pour conclure de façon ferme sur les critères de jugement secondaires (inflation du risque alpha lié à la multiplication des tests).

## Question 12

**Réponses exactes : A – B – D – E**

*Commentaires*

A. Vrai : table 2 aucun intervalle de confiance ne contient la valeur 0.

B. Vrai : inflation du risque alpha (tests multiples).

C. Faux : p1604 colonne 2 : The per-protocol analysis of the secondary outcomes yielded results similar to those of the intention-to-treat analysis, except that there was a significant between group différence in the scores on the EQ-5D visual-analogue scale and *not in the scores on the KOOS symptoms subscale* (Table S3 in the Supplementary Appendix). L'inflation du risque alpha peut expliquer qu'on observe des résultats significatifs à tort alors qu'ici les résultats sont non significatifs.

D. Vrai : l'analyse per protocole est réalisée sur des effectifs plus faibles (25 et 26 sujets) que l'analyse en intention de traiter (50 sujets par groupe) conduisant à une perte de puissance.

E. Vrai : la comparabilité des groupes n'est plus garantie et il pourrait y avoir des facteurs pronostiques distribués différemment dans les deux groupes.

## Question 13

**Réponses exactes : A – E**

*Commentaires*

A. Vrai : limite les biais de classement différentiel sur certains critères de jugement secondaires (temps pour le « timed up-and-go test », temps moyen au test de marche de 20 m, poids, quantité de médicaments antidouleur...).

B. Faux : la standardisation de l'intervention n'intervient pas sur la mesure des critères de jugement.

C. Faux : intervient sur le biais de sélection et de confusion, pas sur les biais de classement.

D. Faux : intervient sur le biais de sélection liés aux patients qui ne suivent pas le traitement de façon optimale.

E. Vrai : condition nécessaire pour que l'aveugle de l'évaluateur puisse être effectif.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 14

**Réponses exactes : B – D**

*Commentaires* Suite à cette étude de bon niveau de preuves malgré certains biais de classement différentiels, on constate une amélioration cliniquement significative dans les deux groupes, et significativement supérieure dans le groupe associant chirurgie puis prise en charge non chirurgicale par rapport au groupe prise en charge non chirurgicale seule, ce qui pourrait faire recommander dans cette population une chirurgie de remplacement prothétique du genou associée à la prise en charge non chirurgicale. Cependant, les effets indésirables statistiquement supérieurs dans le groupe chirurgie puis prise en charge non chirurgicale pondèrent cette recommandation, d'autant plus que la prise en charge non chirurgicale seule a déjà montré son efficacité dans cette population. On ne peut par ailleurs pas recommander un remplacement prothétique seul ou une ou plusieurs composantes de la prise en charge non chirurgicale car elles n'ont pas fait comme tel l'objet de l'expérimentation.

## Question 15

**Réponses exactes : A – C – D**

*Commentaires*

A. Vrai : 2 centres dans le même pays et *a fortiori* des cliniques.

B. Faux : les recommandations sont internationales (p1605 colonne 2 : The combination of nonsurgical treatments that we administered complies with *international recommendations* on the treatment of knee osteoarthritis,[6,7] which increases the generalizability of the results).

C. Vrai : seulement les patients avec des douleurs modérées à sévères (< 60 sur EVA) ont été inclus, pas ceux avec des douleurs très sévères (≥ 60 sur EVA).

D. Vrai : 180 patients incapables de participer au programme de prise en charge non chirurgical et 145 patients incapables de venir au lieu de traitement sur 1475 éligibles » 20 % des patients éligibles : biais de sélection possible.

E. Faux : le fait qu'il n'y ait eu qu'un seul évaluateur joue sur la validité interne de l'étude pas sur la généralisabilité des résultats qui correspond à la validité externe.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 5 Essai d'équivalence

Benoît Marin, Pierre-Marie Preux

| | |
|---|---|
| **Titre de l'article** | **Equivalence of Generic Glatiramer Acetate in Multiple Sclerosis – A Randomized Clinical Trial** |
| **Auteurs** | Cohen J., Belova A., Selmaj K., Wolf C., Sormani M.P., Oberyé J., van den Tweel E., Mulder R., Koper N., Voortman G., Barkhof F. ; Glatiramer Acetate Clinical Trial to Assess Equivalence With Copaxone (GATE) Study Group |
| **Référence** | JAMA Neurol. 2015 Dec ; 72(12):1433-41. |
| **Lien vers l'article** | http://jamanetwork.com/journals/jamaneurology/fullarticle/2451333 |

**1** **Quelle caractéristique s'applique à l'objectif principal de cette étude ?**

a. Étude d'équivalence unilatérale
b. Étude de non-infériorité
c. Étude de bioéquivalence
d. Étude d'équivalence
e. Étude de supériorité

**2** **Quelles sont les caractéristiques applicables à cette étude ?**

a. Étude analytique
b. Étude en trois bras parallèles
c. Étude de recherche biomédicale
d. Essai clinique
e. Étude de phase III

**3** **Quelles propositions sont applicables à la méthode de randomisation ? *Gb : brand glatirameracetate ; Gg : Genericglatirameracetate***

a. Stratifiée sur le centre (hôpital) de recrutement
b. Décentralisée
c. Réalisée via une plateforme internet
d. 3 patients sur 4 recevaient soit le Gb, soit le Gg
e. Stratifiée sur le nombre de lésions gadolinium-positives évalué lors du screening



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **En lien avec l'objectif principal, quelles formulations correspondent d'après vous au résultat principal de cette étude ? *Gb : brand glatirameracetate ; Gg : Genericglatirameracetate***

a. Le ratio (et son intervalle de confiance à 95 %) de l'effet dans le groupe Gg par rapport à l'effet dans le groupe Gb est compris dans la marge d'équivalence définie *a priori*
b. Le degré de significativité de la différence d'effet entre le groupe placebo et le groupe Gg est < 0,001
c. Le degré de significativité de la différence d'effet entre le groupe placebo et le groupe Gb est < 0,001
d. Le degré de significativité de la différence d'effet entre le groupe placebo et le groupe Gb+Gg est < 0,001
e. L'intervalle de confiance à 95 % du ratio de l'effet dans le groupe Gg par rapport à l'effet dans le groupe Gb, en analyse per protocol, est de 0,880-1,368

**5** **Quelle stratégie d'analyse permet d'obtenir la réponse principale dans le cadre d'un essai d'équivalence ? (par exemple Gg vs Gb) ? *Gb : brand glatirameracetate ; Gg : Genericglatirameracetate***

a. Analyse principale basée sur la population en intention de traiter
b. Réalisation d'un test de différence d'effet entre les 2 traitements à comparer et obtention d'un résultat non significatif
c. Calcul de la différence d'effet entre le groupe placebo et le groupe Gg et positionnement de cette différence et intervalle de confiance à 95 % par rapport à la marge d'équivalence
d. Calcul du ratio (et intervalle de confiance à 95 %) de l'effet dans le groupe Gg par rapport à l'effet dans le groupe Gb et positionnement de ce ratio par rapport à la marge d'équivalence définie *a posteriori*
e. Calcul du ratio (et intervalle de confiance à 95 %) de l'effet dans le groupe Gg par rapport à l'effet dans le groupe Gb et positionnement de ce ratio parrapport à la marge d'équivalence définie *a priori*

**6** **Quelle caractérisation peut-on faire du critère de jugement principal ?**

a. Il s'agit d'un critère de jugement « clinique »
b. Il s'agit d'un critère de jugement « composite »
c. Il s'agit d'un critère de jugement « intermédiaire »
d. Il s'agit d'un critère de jugement « paraclinique »
e. Il s'agit d'un critère de jugement combinant efficacité et tolérance

**7** **Quelle stratégie a été appliquée afin d'assurer la validité et fiabilité du critère de jugement principal ?**

a. Évaluation des IRM par des lecteurs en insu
b. Double lecture indépendante des IRM par des radiologues qualifiés et formés
c. Lecture conjointe, simultanée et collégiale des IRM par 3 experts
d. En cas de besoin, intervention d'un 3e expert en insu pour prise d'une décision concernant le critère de jugement principal
e. Mise en place d'un auto-questionnaire auprès des patients



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**8** **Quelles propositions sont exactes concernant la marge d'équivalence du ratio « nombre de lésions nouvelles et persistantes pendant les mois 7 à 9 » entre Gg et Gb ? *Gb : brand glatirameracetate ; Gg : Genericglatirameracetate***

a. La marge d'équivalence a été fixée à 1,750-0,727
b. La valeur maximale du ratio d'efficacité compatible avec la mise en évidence d'une équivalence était de 1,10 environ
c. La borne supérieure de la marge d'équivalence a été fixée à 1,375
d. La borne de non-infériorité supérieure a été fixée à 1,375
e. La borne de non-infériorité inférieure a été fixée à 0,727

**9** **Que pouvez-vous dire concernant le tableau 1 ?**

a. Les caractéristiques présentées entre les trois groupes ne sont pas comparables
b. Il aurait été utile, afin de conclure concernant la comparabilité initiale des groupes, de disposer des résultats de tests statistiques comparant les données entre les groupes
c. Les patients du groupe placebo sont significativement plus jeunes et plus gravement atteints que les sujets inclus dans les autres groupes, ce qui peut être responsable d'un biais de sélection
d. Les groupes obtenus par randomisation sont globalement comparables pour les caractéristiques présentées
e. Les éléments présentés dans ce tableau illustrent le bon déroulement de la randomisation

**10** **Quelles critiques méthodologiques peut-on formuler concernant cette étude ?**

a. Utilisation d'un critère de jugement paraclinique plutôt que clinique
b. Recrutement des patients dans 17 pays (recrutement trop large pour être représentatif)
c. Inclusion d'un bras placebo
d. Les auteurs présentent l'analyse per protocole en soutien de l'analyse en intention de traiter, alors que la pratique dans les essais d'équivalence est de fonder le résultat principal sur l'analyse per protocol (et l'analyse en intention de traiter en soutien)
e. Déséquilibre des bras de randomisation, ce qui réduit la puissance de l'étude

**11** **D'après vous quel est l'intérêt de l'inclusion d'un groupe placebo dans le design de cette étude ? *Gb : brand glatirameracetate ; Gg : Genericglatirameracetate***

a. Le groupe placebo permet d'étudier l'histoire naturelle de la maladie et d'adapter au besoin le nombre de sujets nécessaire en cours d'étude (study sensitivity)
b. Le groupe placebo permet d'évaluer la sensibilité du critère de jugement principal paraclinique par rapport au critère de jugement clinique
c. Il s'agit de vérifier que le traitement générique (Gg) et le traitement fabriqué par le laboratoire (Gb) sont significativement efficaces sur le critère de jugement principal par rapport à un placebo
d. Étant donné que des erreurs lors de la conduite de l'étude peuvent biaiser les résultats d'une étude d'équivalence et aboutir à tort à une conclusion d'équivalence, il convient de vérifier la supériorité de ces produits (Gb et Gg) par rapport au placebo
e. Le groupe placebo n'a aucun intérêt



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**12** **L'inclusion d'un groupe placebo est une préconisation de l'Agence Européenne du Médicament dans le cadre des essais d'équivalence afin de vérifier que les produits expérimentaux procurent bien un bénéfice par rapport au groupe placebo. En effet, une erreur dans le design ou lors de la conduite d'une étude d'équivalence peut aboutir à tort à la conclusion d'une équivalence de stratégies inefficaces. Que pouvez-vous dire concernant (i) les choix méthodologiques menés afin de maîtriser au maximum les problématiques éthiques générés par cette contrainte et (ii) l'impact de cette contrainte sur les résultats de l'étude ?** ***Gb : brand glatirameracetate ; Gg : Generic glatirameracetate***

a. Le ratio de randomisation mis en place permettait qu'un nombre limité de sujets reçoive le placebo
b. Uniquement des patients présentant une maladie peu active ont été inclus
c. Les patients inclus dans l'étude (dont les patients du groupe placebo) pouvaient à la suite de cette étude bénéficier d'un traitement par Gg pendant 15 mois
d. On ne peut pas tenir compte des résultats de cette étude compte tenu de cette problématique éthique
e. Le nombre de retrait de consentement était plus important dans le groupe placebo que dans les deux groupes expérimentaux Gb et Gg

**13** **Quels sont les résultats secondaires/complémentaires apportés par cette étude.** ***Gb : brand glatirameracetate ; Gg : Generic glatirameracetate***

a. L'étude montre clairement l'équivalence significative du Gg par rapport au Gb concernant les critères cliniques
b. Le profil de tolérance sur 9 mois du Gg semble acceptable
c. On observe une réduction significative de la rechute dans les groupes Gb et Gg par rapport au groupe placebo
d. Une supériorité du Gg par rapport au Gb peut être raisonnablement argumentée compte tenu du ratio obtenu
e. On observe une augmentation significative du nombre de patients sans rechute dans les groupes Gb et Gg par rapport au groupe placebo

**14** **Quelles remarques peut-on formuler concernant la forme de cet article**

a. Les références sont présentées selon les recommandations en vigueur
b. Globalement, la structuration classique d'un article scientifique est respectée
c. Contrairement à la règle habituelle en communication scientifique, le titre des figures est situé au-dessus de celles-ci
d. En accord avec la règle habituelle en communication scientifique, le titre des tableaux est situé au-dessus de celles-ci
e. Les informations relatives au nombre de patients considérés pour être inclus (1549) et au nombre de patients randomisés (796) auraient dû être présentés dans les méthodes



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**15** **Selon la HAS, quel est le niveau de preuve de cette étude et à quel grade de recommandations celle-ci peut-elle contribuer**

a. Niveau de preuve 1 (étude de forte puissance, bien conduite)
b. Niveau de preuve 2 (étude d'équivalence)
c. L'étude peut contribuer à la formulation de recommandations de grade B – Présomption scientifique (étude d'équivalence)
d. L'étude peut contribuer à la formulation de recommandations de grade A – Preuve scientifique établie (étude de forte puissance, bien conduite)
e. Niveau de preuve 3 (étude inacceptable éthiquement, de faible puissance en raison des groupes déséquilibrés)



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 1

**Réponse indispensable : D**

*Commentaires*

Gb : brand glatirameracetate ; Gg : Genericglatirameracetate

Il ne s'agit pas d'un essai d'équivalence unilatéral (réponse A fausse) ou de non-infériorité (réponse B fausse), dans ces deux derniers cas on positionnerait le ratio de l'effet d'un traitement par rapport à l'autre traitement vis-à-vis d'une seule borne (de non-infériorité). Or ici ce ratio est positionné vis-à-vis de deux bornes : les marges d'équivalence. Il s'agit clairement d'un essai d'équivalence (réponse D indispensable) du Gg par rapport au Gb. Il ne s'agit pas d'une étude de bioéquivalence (réponse C fausse) qui est destinée à comparer le comportement in vivo d'un médicament à celui du médicament de référence. Même si un bras placebo est inclus dans le design de l'étude pour vérifier la supériorité des deux produits Gb et Gg il ne fallait pas cocher la réponse E car cet aspect ne fait pas partie à proprement parler de l'objectif principal d'équivalence.

## Question 2

**Réponses exactes : B – C – D – E**

*Commentaires* Une étude de type essai clinique (réponse D exacte), et plus globalement une étude d'évaluation de l'efficacité et tolérance d'un nouveau médicament par rapport au traitement de référence (phase III, réponse E exacte) doit être classée réglementairement en recherche biomédicale (réponse C exacte). Il ne s'agit pas d'une étude observationnelle analytique (réponse A fausse). La randomisation aboutit à 3 groupes suivis dans le temps (3 groupes parallèles, réponse B exacte).

## Question 3

**Réponses exactes : C – E**

*Commentaires* La randomisation est stratifiée sur le nombre de lésions gadolinium-positives évalué lors du screening (voir paragraphe « Randomization and Masking », réponse E exacte). Elle n'est pas stratifiée sur l'hôpital mais la zone géographique (réponse A fausse). Elle est centralisée (réponse B fausse) et opérée par une plateforme internet couplée à un système vocal (réponse C exacte). Le ratio de randomisation est 4,3 : 4,3:1 pour Gb, Gg et placebo, ainsi 8,6 patients pour 9,6 patients reçoivent soit le Gb soit le Gg (86,6 %) ce qui ne correspond pas à 75 % (réponse D fausse).

## Question 4

**Réponse exacte : E**

**Indispensable : A**

*Commentaires* La démarche de prise de décision dans le cadre d'un essai d'équivalence consiste à placer le ratio de l'effet dans le groupe A par rapport à l'effet dans le groupe B (ou la différence d'effet) ainsi que son intervalle de confiance à 95 % par rapport à la marge d'équivalence définie à priori (réponse A indispensable). L'analyse de référence est l'analyse per protocol (réponse E exacte) même si dans l'article, les auteurs mettent l'analyse en intention de traiter en avant. Les résultats relatifs aux comparaisons Gb, Gg, Gg+Gb par rapport au placebo n'interviennent pas dans le cadre de l'objectif principal (réponses B, C, D fausses).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 5

**Réponse exacte : E**

**Indispensable : E**

**Inadmissible : A**

*Commentaires* L'analyse principale d'une étude d'équivalence est l'analyse per protocol et non l'analyse en intention de traiter (réponse A inadmissible). La démarche de prise de décision dans le cadre d'un essai d'équivalence consiste à placer le ratio de l'effet dans le groupe A par rapport à l'effet dans le groupe B (ou la différence d'effet) ainsi que son intervalle de confiance à 95 % par rapport à la marge d'équivalence définie à priori (réponse E indispensable, réponse D (marge « *a posteriori* ») fausse). Un test de différence non significatif ne saurait permettre d'aboutir à un jugement d'équivalence (réponse B fausse). De même, la différence d'effet entre le groupe placebo et le groupe Gg ne dit rien concernant l'équivalence Gg/Gb (réponse C fausse).

### Question 6

**Réponses exactes : C – D**

*Commentaires* Le critère de jugement principal est le nombre de lésions nouvelles et persistantes à l'IRM pendant les mois 7 à 9. Il ne s'agit pas d'un critère clinique (exemple : rémission clinique ; rechute, réponse A fausse) mais d'un critère paraclinique (réponse D exacte) qui est « intermédiaire », c'est-à-dire que sa modification par le traitement doit aboutir à une modification de l'état clinique du patient (réponse C exacte). Ce point est par ailleurs discuté par les auteurs dans la discussion (limite/avantage de ce choix méthodologique). Ce critère n'est pas composite (il n'est pas composé de différentes entités, réponse B fausse) et ne combine pas efficacité et tolérance (réponse E fausse).

### Question 7

**Réponses exactes : A – B – D**

*Commentaires* Dans « Procédures » il est indiqué que 2 radiologues qualifiés et formés évaluent indépendamment et en insu les IRM (réponses A et B exactes). Un 3e expert intervient en cas de besoin pour la décision concernant le critère de jugement (réponse D exacte, réponse C fausse). Le questionnaire patient mis en place concerne un critère de jugement secondaire (tolérance, réponse E fausse).

### Question 8

**Réponses exactes : B – C**

*Commentaires* Dans « statistical analysis », 1re colonne page 1436 il est indiqué que la marge d'équivalence est fixée à 0,727-1,375 (réponse C exacte, réponse A fausse). Il est aussi indiqué qu'en conséquence, la valeur maximale du ratio d'efficacité compatible avec la mise en évidence d'une équivalence était de 1,10 environ : « the maximal allowable différence between the point estimates to show equivalence would be approximately 10 % » (réponse B exacte). Des bornes dites de non-infériorité « supérieure » ou « inférieure » n'existent pas (réponses D et E fausses).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 9

**Réponses exactes : D – E**

*Commentaires* Le tableau 1 présente, comme cela est attendu dans une publication d'essai clinique, les caractéristiques initiales des sujets inclus et randomisés. C'est l'examen des indices de position et de dispersion par un lecteur muni de pertinence clinique qui doit permettre de déterminer si les groupes sont comparables ou dissemblables. Un test statistique comparant les données entre les groupes n'aurait pas de sens ici (on testerait le hasard, puisque c'est la randomisation qui a permis de constituer les groupes) (réponse B fausse). Globalement les sujets présentent des caractéristiques proches (comparabilité vérifiée, réponses A et C fausses, réponse D exacte). La randomisation menée correctement a joué son rôle (réponse E exacte).

### Question 10

**Réponses exactes : A – D**

*Commentaires* Le choix d'utiliser un critère paraclinique est argumenté par les auteurs dans la discussion mais reste une limite de l'étude (réponse A exacte). Le recrutement très multicentrique (17 pays) favorise la généralisation des résultats (réponse B fausse). L'inclusion d'un bras placebo n'est pas une limite et en l'occurrence une préconisation de l'Agence Européenne du Médicament pour les essais d'équivalence de ce type (réponse C fausse). Les auteurs présentent principalement l'analyse en intention de traiter pour argumenter l'équivalence (et l'analyse per protocol en soutien) alors que dans ce type d'étude c'est bien l'analyse per protocol qui est la plus importante (réponse D exacte). Le déséquilibre des groupes de randomisation n'est en rien critiquable, au contraire, cela permet de limiter le nombre de sujets recevant le placebo (réponse E fausse).

### Question 11

**Réponses exactes : C – D**

**Inadmissible : E**

*Commentaires* Le groupe placebo permet de vérifier que le traitement générique (Gg) et le traitement fabriqué par le laboratoire (Gb) sont significativement efficaces sur le critère de jugement principal par rapport à un placebo. En effet une étude de ce type avec des erreurs dans la planification et la conduite de l'étude peut aboutir à tort à un jugement d'équivalence. Ce point est argumenté page 1436, 1re colonne (réponses C et D exactes, réponse E inadmissible). Il ne s'agit pas d'évaluer l'histoire naturelle pour réévaluer le nombre de sujets nécessaires (réponse A fausse), ni de tester la sensibilité du critère paraclinique par rapport au critère clinique (réponse B fausse).

### Question 12

**Réponses exactes : A – B – C**

*Commentaires* Dans un tel contexte critiquable éthiquement (inclusion d'un groupe placebo alors qu'un traitement actif existe), les auteurs ont utilisé un ratio de randomisation déséquilibré pour limiter le nombre de sujets randomisés dans le groupe placebo (réponse A exacte), seuls les patients présentant une maladie peu active ont été inclus (réponse B exacte), tous les patients peuvent à l'issue de l'essai, bénéficier d'un traitement par Gg pendant 15 mois (réponse C exacte). Le nombre de retraits est resté faible dans le groupe placebo (réponse E fausse). Cette étude respecte les recommandations de l'Agence Européenne du Médicament (réponse D fausse).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 13**

**Réponse exacte : B**

**Indispensable : B**

**Inadmissible : D**

*Commentaires* Le tableau 3 permet d'identifier que le profil de tolérance du Gg ne pose pas de problème par rapport au Gb (réponse B indispensable). On ne met pas en évidence d'équivalence du Gg par rapport au Gb sur les critères cliniques (réponse A fausse, l'étude n'était par ailleurs pas dimensionnée pour cela). De même, on ne met pas en évidence de réduction significative de rechutes ou augmentation significative du nombre de patients sans rechute dans les groupes Gg et Gb par rapport au placebo, l'étude n'a pas été dimensionnée pour cela, et cela n'a pas été testé (réponse C et E fausses). Le ratio obtenu ne permet pas d'argumenter la supériorité du Gg (réponse D inadmissible).

**Question 14**

**Réponses exactes : A – B – C – D**

*Commentaires* Les références sont bien présentées selon les recommandations en vigueur (ici Vancouver) : auteurs. titre de l'article. nom de la revue. année ; volume : pages (réponse A exacte). La structure Introduction/Méthodes/Résultats et Discussion est bien respectée (réponse B exacte). Habituellement on place le titre des figures sous celles-ci (réponse C exacte), et le titre des tableaux au-dessus de ceux-ci (réponse D exacte). Les nombre de sujets éligibles et inclus font partie des résultats (réponse E fausse).

**Question 15**

**Réponses exactes : A – D**

*Commentaires* Il s'agit d'une étude expérimentale de type essai clinique d'équivalence bien conduit et de forte puissance, aboutissant logiquement au niveau de preuve 1 et grade de recommandations A (réponse A et D indispensables, réponses B, C E fausses).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 6 Essai de non-infériorité

Agnès Dechartres

| | |
|---|---|
| **Titre de l'article** | **Outpatient *versus* inpatient uterine polyp treatment for abnormal uterine bleeding : randomised controlled non-inferiority study** |
| **Auteurs** | Cooper NA, Clark TJ, Middleton L, Diwakar L, Smith P, Denny E, Roberts T, Stobert L, Jowett S, Daniels J ; OPT Trial Collaborative Group. |
| **Référence** | BMJ 2015 ; 350:h1398 |
| **Lien vers l'article** | **http://www.bmj.com/content/350/bmj.h1398** |

**1** **Parmi les propositions suivantes, laquelle ou lesquelles correspondent à l'objectif principal de cette étude ?**

a. Évaluer si l'ablation de polypes utérins (polypectomie) en ambulatoire est faisable en pratique courante
b. Évaluer la proportion de centres pratiquant l'ablation de polypes utérins (polypectomie) en ambulatoire au Royaume-Uni
c. Démontrer que l'ablation de polypes utérins (polypectomie) en ambulatoire est au moins aussi efficace que la polypectomie sous anesthésie générale sur les saignements à 6 mois
d. Évaluer les facteurs associés à la pratique de l'ablation de polypes utérins (polypectomie) en ambulatoire au Royaume-Uni
e. Monter que l'ablation de polypes utérins (polypectomie) en ambulatoire permet de limiter les coûts par rapport à la polypectomie sous anesthésie générale

**2** **Quel(s) est/ sont le ou les avantage(s) du traitement en ambulatoire par rapport à la prise en charge habituelle mis en avant par les auteurs pour justifier le choix d'un essai de non-infériorité ?**

a. Le risque de saignement per et postopératoire est moins important avec le traitement ambulatoire
b. Le traitement ambulatoire se fait sous anesthésie locale
c. Le traitement ambulatoire peut se faire au moment de l'hystéroscopie diagnostique
d. Le traitement ambulatoire est moins douloureux que le traitement chirurgical sous anesthésie générale
e. Le traitement ambulatoire pourrait permettre de limiter les hospitalisations



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**3** **Parmi les mots-clés suivants, quel est ou quels sont ceux qui correspondent au plan expérimental de cette étude ?**

a. Prospectif
b. Expérimental
c. Ouvert
d. Équivalence
e. Longitudinal

**4** **Quels sont les éléments rapportés dans cette étude qui sont en faveur d'un schéma d'étude pragmatique ?**

a. La randomisation par minimisation
b. Les critères d'inclusion assez larges
c. Le fait d'avoir différents types de centres notamment ceux proposant un service « see and treat »
d. La présence d'un comité indépendant monitorant les données
e. Le fait que les médecins puissent choisir la technique de polypectomie

**5** **Parmi les propositions suivantes concernant la méthode de randomisation, laquelle ou lesquelles sont vraies ?**

a. La randomisation était équilibrée
b. La randomisation était stratifiée
c. La randomisation était centralisée
d. La randomisation était faite par minimisation
e. La randomisation était en cluster

**6** **Le critère de jugement principal était**

a. L'efficacité du traitement sur les saignements évaluée par le médecin
b. Un critère composite
c. Une variable qualitative
d. À risque de biais de classement
e. De nature subjective

**7** **Dans cette étude, la marge de non-infériorité était**

a. La différence minimale que les auteurs souhaitaient mettre en évidence entre les deux groupes
b. Justifiée par les auteurs par des avantages en termes de coût
c. Justifiée par les auteurs par des avantages en termes de commodité pour les femmes
d. Justifiée par les capacités de recrutement des centres
e. Utilisée pour estimer le nombre de patientes à inclure



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**8** **Selon les auteurs, cette étude a utilisé pour l'analyse du critère de jugement principal**

a. Une marge de non-infériorité de 25 % (différence absolue)
b. Une analyse en intention de traiter
c. Une analyse per protocole
d. Des risques relatifs bruts
e. Des analyses de type mesures répétées ajustées sur la valeur à baseline comme indiqué dans la Figure 2

**9** **Dans cette étude, la population d'analyse pour l'analyse principale était**

a. L'ensemble des femmes consultant dans l'un des 31 centres de l'essai pour un saignement utérin anormal avec un polype diagnostiqué à l'hystéroscopie pour lesquelles la polypectomie était faisable en ambulatoire, chez qui on ne suspectait pas de cancer et qui ne devaient pas subir une intervention chirurgicale
b. L'ensemble des femmes randomisées selon leur groupe de randomisation initial conformément au principe de l'intention de traiter
c. L'ensemble des femmes randomisées selon leur groupe de randomisation initial ayant une visite de suivi à 6 mois
d. L'ensemble des femmes randomisées ayant reçu l'intervention allouée
e. L'ensemble des femmes randomisées selon leur groupe de randomisation initial ayant été suivies jusqu'à la fin de l'essai

**10** **Parmi les éléments suivants concernant le tableau 1, quel(s) est (sont) celui ou ceux qui sont vrais ?**

a. Il permet de conclure à une comparabilité initiale des groupes
b. En l'absence des p et/ ou intervalles de confiance à 95 %, il n'est possible de se prononcer
c. La proportion de polypes glandulaires dans le groupe traitement ambulatoire était non inférieure à la proportion dans le groupe traitement sous anesthésie générale
d. La proportion de femmes avec un saignement post ménopause était équivalente entre les deux groupes
e. L'âge aurait dû être décrit avec la médiane et non avec la moyenne

**11** **Concernant le résultat principal de cette étude**

a. La différence absolue de risque entre les deux groupes était 0,91 (IC 95 % 0,82 à 1,02)
b. Il y avait une différence relative de 7 % sur le critère de jugement principal entre les deux groupes
c. 73 % des femmes du groupe traitement sous anesthésie générale ont rapporté une efficacité sur les saignements à 6 mois
d. La différence entre les deux groupes était exprimée par un risque relatif
e. L'analyse en intention de traiter et l'analyse per protocol retrouvaient des résultats cohérents



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**12** **Il est possible de conclure à la non-infériorité de la polypectomie en ambulatoire par rapport à la polypectomie sous anesthésie générale car**

a. La valeur du risque relatif en intention de traiter était de 0,91 donc > 0,75
b. Toutes les valeurs de risque relatif étaient > 0,25
c. La borne inférieure de l'intervalle de confiance à 95 % du risque relatif était > 0,75
d. La borne inférieure de l'intervalle de confiance à 95 % du risque relatif était > 0,25
e. L'intervalle de confiance à 95 % du risque relatif contenait la valeur 1

**13** **Le tableau A3 présente le résultat d'analyses en sous-groupes. Que peut-on dire de ces analyses ?**

a. Les auteurs rapportent les avoir planifiées à l'avance
b. Les effectifs étaient relativement équilibrés au sein de chaque sous-groupe grâce à la minimisation
c. Les résultats des tests d'interaction peuvent être expliqués par le choix du risque de première espèce
d. Le risque relatif ne variait pas significativement selon le type de polype ou leur localisation
e. Ces analyses ne peuvent être pas être interprétées ici car le résultat principal n'est pas significatif

**14** **Quel(s) élément(s) limite(nt) la généralisation des résultats de cette étude ?**

a. La méthode de randomisation utilisée
b. Le design pragmatique de l'essai
c. Le fait que l'essai soit multicentrique
d. Le fait que l'avantage énoncé en termes de coût plus faible pour les femmes ne soit pas forcément valable dans d'autres contextes de soins
e. L'absence d'aveugle

**15** **Quelle(s) est (sont) la ou les implications potentielles de cette étude ?**

a. Il faut abandonner la polypectomie utérine sous anesthésie générale afin de limiter des hospitalisations inutiles
b. Il faut privilégier la polypectomie utérine en ambulatoire car cette intervention a montré qu'elle était non-inférieure à la polypectomie sous anesthésie générale et qu'elle présentait des avantages notamment en termes de coût
c. Il n'y a pas lieu de proposer une polypectomie en ambulatoire en raison de la plus grande difficulté à enlever des polypes et de l'acceptabilité qui était moindre qu'avec une intervention sous anesthésie générale
d. Les deux interventions peuvent être envisagées mais il faut clairement informer les femmes sur les avantages et inconvénients de chaque type d'intervention
e. La prise en charge en ambulatoire peut être généralisée à d'autres spécialités chirurgicales



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 1**

**Réponse exacte : C**

**Inadmissibles : B – D**

*Commentaires* Cette étude est un essai contrôlé randomisé. La question porte ici sur l'objectif principal et doit donc correspondre au critère de jugement principal c'est-à-dire montrer que la polypectomie ambulatoire est non-inférieure (c'est-à-dire au moins aussi efficace) en termes d'efficacité sur les saignements à la polypectomie sous anesthésie générale. Concernant l'item A, il ne s'agit pas ici d'une étude de faisabilité. Les auteurs comparent l'acceptabilité de la procédure et la satisfaction des patientes entre les deux types d'intervention mais il s'agit de critères secondaires. Les items B et D sont faux, ils ne correspondent pas à des objectifs d'évaluation thérapeutique. Il n'y a pas d'évaluation des coûts dans cette étude (item E faux).

**Question 2**

**Réponses exactes : B – C – E**

*Commentaires* Il est licite de faire un essai de non-infériorité quand le traitement que l'on souhaite évaluer présente des avantages autres que l'efficacité par rapport au traitement habituellement utilisé. Dans cette situation, il est nécessaire de montrer 1) que le nouveau traitement n'est pas inférieur au traitement habituellement utilisé en termes d'efficacité et 2) que le traitement apporte par exemple un meilleur profil de tolérance, un coût inférieur, une meilleure acceptabilité des patients (par exemple prise de comprimés par rapport à des injections).

Dans cette étude, la polypectomie en ambulatoire permet de ne pas avoir recours à une anesthésie générale, l'intervention se faisant sous anesthésie locale, de ne pas hospitaliser les patientes et de faire dans certains centres le geste directement au moment du diagnostic.

**Question 3**

**Réponses exactes : A – B – C – E**

**Indispensables : A – B**

*Commentaires* Il s'agit ici d'un essai contrôlé randomisé en ouvert (car l'aveugle n'est pas possible), de non-infériorité et non pas d'équivalence. Dans un essai d'équivalence, on n'autorise le nouveau traitement ni à être inférieur, ni à être supérieur, alors que dans l'essai de non-infériorité, le nouveau traitement peut être supérieur.

Un essai contrôlé randomisé est une étude expérimentale et est forcément longitudinal (on suit le patient dans le temps) et prospectif.

**Question 4**

**Réponses exactes : B – C – E**

*Commentaires* L'aspect pragmatique signifie que l'on est plus proche de la pratique réelle avec de la variabilité dans les patients traités et dans l'administration de l'intervention. Les essais pragmatiques s'opposent aux essais explicatifs dans lesquels on cherche à maximiser l'efficacité du traitement. Ces essais explicatifs ont fréquemment des critères d'inclusion très restrictifs, l'administration du traitement est très standardisée et vérifiée. En réalité, la notion d'explicatif/pragmatique n'est pas vraiment une notion binaire : il y a un continuum entre ces deux notions.

Dans cette étude, le fait d'avoir des critères d'inclusion assez larges (seules les femmes pour lesquelles la polypectomie en ambulatoire ne semblait pas faisable, celles avec une suspicion de cancer et celles nécessitant une autre intervention chirurgicale étaient exclues), de laisser



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

le choix aux médecins sur la manière de réaliser la polypectomie et ne pas standardiser l'intervention, l'inclusion de centres avec différentes pratiques sont des éléments en faveur du côté pragmatique.

La randomisation par minimisation est un élément en faveur d'une bonne validité interne et non externe. La présence d'un comité indépendant monitorant les données est un élément important pour la sécurité des patients dans l'étude.

### Question 5

**Réponses exactes : A – C – D**

*Commentaires* Il y a deux éléments essentiels à évaluer concernant la randomisation : la méthode pour générer la séquence et l'assignation secrète ou clause d'ignorance (« allocation concealment » en anglais). Ici la méthode pour générer la séquence est une randomisation par minimisation pour limiter les déséquilibres sur les éléments suivants : le type de saignement, la localisation et le type de polype. Cette technique consiste à équilibrer les groupes à chaque nouvelle inclusion Le groupe est déterminé pour chaque patient par un algorithme prédéfini tenant compte de ses caractéristiques afin de minimiser les différences entre les groupes. L'allocation par minimisation est souvent utilisée lorsque le nombre de sujets nécessaire est relativement faible et si on a trop de variables sur lesquelles on voudrait stratifier. Cette technique a tendance à être de plus en plus utilisée dès qu'une randomisation centralisée est possible, car elle est particulièrement efficace pour équilibrer les groupes.

La randomisation était équilibrée car on randomisait autant de patients dans le bras expérimental que dans le bras contrôle.

Concernant le principe d'assignation secrète, il est respecté ici car la randomisation est centralisée via internet.

Il s'agit ici d'une randomisation individuelle et non en cluster. Cette dernière ne s'applique que lorsque la randomisation concerne des groupes de patients et non directement des individus.

### Question 6

**Réponses exactes : B – C – D – E**

*Commentaires* Le critère de jugement principal est ici l'efficacité sur les saignements déterminée à partir de l'évaluation par les patientes de leurs saignements à 6 mois. Il s'agit un critère rapporté par le patient et non pas évalué par le médecin. Le critère était un critère binaire (succès ou échec) donc qualitatif. Ce critère est basé sur l'évaluation des patientes de l'intensité de leurs saignements, critère qui est très subjectif, en ouvert (car l'aveugle n'était pas possible) donc le risque de biais de classement est important. L'évaluation n'utilisait pas les mêmes critères selon le type de saignement mais cela n'est pas pour autant un critère composite. Un exemple de critère composite aurait été saignement ou infection à 6 mois.

### Question 7

**Réponses exactes : B – C – E**

*Commentaires* La marge de non-infériorité est un élément crucial dans un tel essai. Elle doit être définie au moment de la planification de l'étude sur des critères de pertinence et d'acceptabilité (jusqu'à quelle limite on considère que la nouvelle intervention est non-inférieure à l'intervention de référence). Plus la marge est petite, plus il faudra inclure de patients. Dans cette étude,



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

la marge semble très large. Le choix d'une marge aussi large est justifié par les auteurs dans le calcul d'effectif par le fait qu'une efficacité un peu moindre serait contre-balancée par des avantages importants en termes de coût et de commodité pour les femmes.

### Question 8

**Réponses exactes : B – C – D**

*Commentaires* Dans un essai de non-infériorité, il est essentiel de définir *a priori* la marge de non-infériorité qui était fixée dans cette étude à 25 % en différence relative et non en différence absolue (la raison pour laquelle la proposition A est fausse). Dans un essai de non-infériorité, on doit faire une analyse en intention de traiter et une analyse per protocole. C'est le seul type d'essai où l'analyse en per protocole est licite car c'est l'analyse la plus conservatrice. En effet, elle augmente les différences et ainsi diminue le risque de conclure à tort. L'analyse en intention de traiter est la règle dans un essai de supériorité car elle permet de limiter les biais liés à l'exclusion des patients de l'analyse et permet de maintenir la comparabilité des groupes issus de la randomisation lors de l'analyse (si le suivi est bien conduit). Elle est en théorie plus conservatrice c'est-à-dire qu'elle devrait tendre à diminuer les différences entre les 2 groupes. Dans un essai de non-infériorité, diminuer les différences entre les 2 groupes aboutirait à conclure plus facilement à la non-infériorité d'où l'importance de l'analyse per protocole.

Dans cette étude, l'effet traitement est estimé par des risques relatifs (ou risk ratios) bruts c'est-à-dire non ajustés.

La réponse E est fausse car le critère de jugement principal était évalué à 6 mois et ne prenait donc pas en compte les mesures répétées. Cette analyse a été faite pour les critères de jugement secondaires.

### Question 9

**Réponse exacte : C**

*Commentaires* La population d'analyse correspond à la population qui a été effectivement prise en compte dans l'analyse. Il ne s'agit donc pas de la population éligible (item A). Ici, les auteurs précisent dans la partie méthodes qu'ils ont fait une analyse en intention de traiter pour l'analyse principale (l'analyse per protocole est définie comme une analyse de sensibilité). Seulement, le flow chart ainsi que le dénominateur dans les analyses montrent que ce n'est pas le cas et qu'ils ont uniquement considéré les femmes ayant une évaluation du critère de jugement à 6 mois. Il y a une contradiction entre le flowchart et le tableau 1 en termes d'effectifs, mais la D est indubitablement fausse car 26 et 39 patientes sont sorties d'étude et ne sont pas analysées.

Une analyse complémentaire aurait pu être faite en remplaçant les données manquantes à 6 mois, en remplaçant ces données par la valeur moyenne du groupe ou sur la base d'une hypothèse de type biais maximum considérant les patientes du groupe test en échec et les patientes du groupe de référence en succès.

### Question 10

**Réponse exacte : A**

**Inadmissible : C**

*Commentaires* Le tableau 1 permet d'évaluer la comparabilité des groupes sur les données à baseline. Cette comparaison doit s'évaluer à l'œil (en vérifiant que les déséquilibres ne sont pas trop importants compte tenu de l'effectif recruté et ne vont pas toujours dans le même sens). Les tests statistiques sont inutiles voire inappropriés (item B faux) car si la randomisation est bien faite, une différence significative serait vraisemblablement due au hasard. Comme il y a



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

fréquemment beaucoup de caractéristiques évaluées dans ce tableau, le risque alpha est très important. De plus, ce sont des tests de supériorité et une absence de différence statistiquement significative ne signifie pas que les deux groupes sont équivalents.

La moyenne est un paramètre approprié pour décrire l'âge qui a généralement une distribution normale (item E faux).

### Question 11

**Réponses exactes : D – E**

**Inadmissible : A**

*Commentaires* Le résultat principal de cette étude correspond au résultat pour le critère de jugement principal. 73 % des femmes (166/228) du groupe traitement ambulatoire et 80 % (168/211) des femmes du groupe traitement sous anesthésie générale rapportaient une efficacité sur les saignements soit un risque relatif en intention de traiter de 0,91 (IC à 95 % 0,82-1,02) et en per protocol à 0,92 (IC à 95 % 0,82-1,02). La réponse A était donc fausse. Les résultats en ITT et en PP étaient cohérents. ; la réponse E était donc vraie. La différence absolue entre les deux groupes était de 7 % ; la réponse B était donc fausse.

### Question 12

**Réponse exacte : C**

**Inadmissible : E**

*Commentaires* Dans un essai de non-infériorité, on doit fixer une marge de non-infériorité. Dans cet essai, elle était de 25 % (différence relative) ce qui correspond à un risque relatif de 0,75. Pour pouvoir conclure à la non-infériorité, il faut que la borne inférieure du risque relatif soit strictement supérieure à 0,75.

### Question 13

**Réponses exactes : A – B – D**

*Commentaires* Les auteurs rapportent dans la partie analyse statistique des méthodes que les analyses en sous-groupe étaient pré-spécifiées (item A vrai). Quand on prévoit des analyses en sous-groupe, il est important d'avoir des effectifs équilibrés au sein de chaque sous-groupe afin de pouvoir interpréter les résultats correctement. La randomisation par minimisation prenait en compte ces critères donc on s'attend à avoir des effectifs relativement équilibrés pour ces analyses ce qui est le cas (item B vrai). Dans ce tableau, aucun test d'interaction n'est significatif ($p > 0,05$) donc les résultats ne peuvent pas être affectés par le risque de première espèce (risque de conclure à tort à une différence qui n'existe pas) (item C faux).

Dans un essai de supériorité, les analyses en sous-groupe peuvent être interprétées à titre exploratoire quand elles sont planifiées à l'avance, en nombre limité et si le résultat principal est significatif (il n'est pas possible de conclure à un bénéfice dans un sous-groupe particulier si le résultat principal est non-significatif). Ici, on est dans le cadre d'un essai de non-infériorité donc l'item E est faux.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 14**

**Réponse exacte : D**

**Inadmissible : C**

*Commentaires* La généralisabilité ou validité externe est le fait que l'on puisse extrapoler les résultats d'une étude à sa pratique. Pour l'évaluer, il faut considérer les caractéristiques des patients, des centres, du contexte de soins. Les résultats d'un essai pragmatique sont plus facilement généralisables que dans un essai explicatif car l'objectif de l'essai pragmatique est de s'approcher de la « vraie vie ». Les résultats d'un essai multicentrique sont plus facilement généralisables que ceux d'un essai monocentrique. La méthode de randomisation et l'aveugle sont des éléments liés à la validité interne. Les éléments de coûts recueillis dans un pays avec un système de protection sociale et économique donné ne sont pas forcément extrapolables dans un autre contexte.

**Question 15**

**Réponse exacte : D**

**Inadmissible : A**

*Commentaires* Cette étude a montré la non-infériorité de la polypectomie en ambulatoire par rapport à la polypectomie sous anesthésie générale mais avec une marge assez large. Les deux interventions sont donc possibles mais il semblerait (même si ce n'est pas le critère de jugement principal), qu'il y ait plus de difficultés à enlever les polypes et que l'acceptabilité soit moindre. La conclusion la plus appropriée est donc d'informer les femmes sur les avantages et inconvénients de chaque type d'intervention afin qu'elles prennent avec leur médecin la décision la plus appropriée à leur situation.

La généralisation du traitement ambulatoire à d'autres spécialités chirurgicales n'est pas possible directement. Il faudrait évaluer l'intérêt du traitement ambulatoire spécialité par spécialité dans des essais contrôlés randomisés.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 7 Essai de non-infériorité

Virginie Gardette, Adeline Gallini, Matthieu Wargny

| | |
|---|---|
| **Titre de l'article** | **Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes** |
| **Auteurs** | Zinman b, Wanner c, Lachin JM, Fitchett D, Bluhmki E, Hantel S, Mattheus M, Devins T, Johansen OE, Woerle HJ, Broedl UC, Inzucchi SE, for the EMPA-REG OUTCOME Investigators |
| **Référence** | N Engl J Med 2015 ; 373:2117-28. |
| **Lien vers l'article** | http://www.nejm.org/doi/full/10.1056/NEJMoa1504720#t = article |

**1** **Indiquez laquelle de ces affirmations correspond à l'hypothèse principale de l'étude, sous les hypothèses précisées dans la partie statistique de l'article**

a. Comparativement au placebo, l'empagliflozine permet de diminuer d'au moins 30 % le risque de décès d'origine cardiovasculaire, IDM ou AVC chez le sujet diabétique de type 2 à haut risque cardiovasculaire
b. Comparativement au placebo, l'empagliflozine n'augmente pas de plus de 30 % le risque de décès d'origine cardiovasculaire, IDM ou AVC chez le sujet diabétique de type 2 à haut risque cardiovasculaire
c. Comparativement au placebo, l'empagliflozine 10 mg diminue d'au moins 30 % le risque de décès d'origine cardiovasculaire, IDM ou AVC chez le sujet diabétique de type 2 à haut risque cardiovasculaire
d. Comparativement au placebo, l'empagliflozine 10 mg n'augmente pas de plus de 30 % le risque de décès d'origine cardiovasculaire, IDM ou AVC chez le sujet diabétique de type 2 à haut risque cardiovasculaire
e. Comparativement au placebo, l'empagliflozine 25 mg diminue d'au moins 30 % le risque de décès cardiovasculaires, IDM ou AVC chez le sujet diabétique de type 2 à haut risque cardiovasculaire

**2** **Concernant le schéma d'étude, il s'agit d'un essai**

a. D'intervention
b. En groupes parallèles
c. En plan factoriel
d. Ouvert
e. De non-infériorité



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**3** **À propos de la méthode, indiquez quelle est (quelles sont) la (les) réponse(s) exacte(s) ?**

a. Le nombre de sujets à inclure d'après les hypothèses de l'étude est de 691
b. Les rendez-vous prédéfinis minimisent le biais de suivi entre les groupes
c. Le seuil choisi pour le risque de première espèce est de 0,05
d. La déclaration d'Helsinki précise le cadre général des recommandations de bonne pratique méthodologique de conduite d'un essai clinique
e. Les analyses statistiques ont été réalisées en aveugle de l'intervention reçue

**4** **À propos des interventions comparées dans cette étude, indiquez quelle est (quelles sont) la (les) réponse(s) exacte(s) ?**

a. L'empagliflozine est comparée à la prise en charge usuelle
b. L'empagliflozine est prescrite à dose fixe
c. Suite à la période de run-in, l'empagliflozine est proposée aux sujets répondeurs à ce médicament
d. La combinaison des résultats des 2 groupes traités par empagliflozine correspond à une méta-analyse
e. Il n'est pas éthique de proposer à un groupe de ne pas recevoir l'empagliflozine car son intérêt a déjà été démontré dans d'autres essais cliniques

**5** **À propos de la randomisation, indiquez quelle est (quelles sont) la (les) réponse(s) exacte(s) ?**

a. Elle prévoit des effectifs identiques d'individus recevant le placebo et l'empagliflozine
b. Elle permet d'assurer la comparabilité des groupes tout au long de l'essai
c. La stratification réalisée dans la randomisation garantit l'absence de biais de sélection
d. Son effectivité peut être vérifiée par la comparaison des caractéristiques des sujets analysés dans l'étude
e. Sa réalisation par informatique est un critère de bonne qualité

**6** **À propos des critères de jugement retenus dans cette étude, indiquez quelle est (quelles sont) la (les) réponse(s) exacte(s) ?**

a. Un critère de qualité d'un critère composite consiste à regrouper plusieurs évènements représentant une entité cliniquement pertinente
b. Le choix d'un critère composite peut être justifié face à des évènements rares
c. Il est possible que seules quelques composantes du critère composite soient en faveur de la différence entre les groupes
d. Lors de l'interprétation de l'effet mesuré par un critère composite, il faut s'intéresser à la fréquence de chaque composante du critère composite
e. Les mesures d'association entre l'exposition et les différents critères inclus dans le critère de jugement composite doivent aller dans le même sens pour que l'effet sur le critère de jugement composite soit interprétable.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**7** **Dans cette étude, à propos des comités statuant sur la survenue ou non des évènements, indiquez quelle est (quelles sont) la (les) réponse(s) exacte(s) ?**

a. Ils permettent de limiter le biais de sélection
b. Ils permettent de limiter le biais de classement
c. Ils permettent de limiter le biais de confusion
d. Ils n'ont aucun impact sur les biais éventuels
e. L'existence de 2 comités distincts augmente le risque de biais de mesure

**8** **À propos de l'analyse statistique retenue dans cette étude, indiquez quelle est (quelles sont) la (les) réponse(s) exacte(s) ?**

a. L'analyse du critère de jugement principal est réalisée en intention de traiter
b. Prendre en compte la mortalité comme un risque compétitif consiste à ne plus comptabiliser les patients décédés comme étant à risque de présenter l'évènement d'intérêt dès leur entrée dans l'étude
c. Le moment de la censure pour l'analyse principale ne correspond qu'à 3 moments possibles : la survenue du critère de jugement principal, de la perte de vue, ou de la date de point de l'essai
d. Tout évènement survenu après la date de censure était comptabilisé
e. Une analyse intermédiaire ayant été réalisée, le risque alpha a été corrigé pour l'analyse finale

**9** **À propos du tableau 1, indiquez quelle(s) affirmation(s) est (sont) correcte(s) ?**

a. Dans le groupe Placebo, 43,9 patients-années ont rencontré le critère composite principal (décès pour maladie cardiovasculaire, infarctus du myocarde non fatal ou accident cardiovasculaire (AVC) non fatal), contre 37,4 patients-années dans le groupe Empagliflozine
b. Le risque de présenter un AVC non fatal était significativement plus important dans le groupe Empagliflozine que dans le groupe Placebo
c. Le risque d'être hospitalisé pour une insuffisance cardiaque était significativement plus important dans le groupe Empagliflozine que dans le groupe Placebo
d. Dans le groupe Empagliflozine, 38 patients sur 4 687 ont présenté une ischémie myocardique silencieuse au cours du suivi
e. Le risque de décéder au cours du suivi, qu'elle qu'en soit la cause, au cours du suivi était significativement plus faible dans le groupe Empagliflozine que dans le groupe Placebo, et estimé ici à une réduction du risque de 32 % ($IC_{95\ \%}$, [-43 % ; -18 %], P < 0,001)



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**10** **À propos de la figure 1 page 2121, indiquez quelle(s) affirmation(s) est (sont) correcte(s) ?**

a. La figure 1.A représente l'évolution de l'incidence cumulée du critère de jugement principal dans les groupes Empagliflozine et Placebo, au cours des 4 années de suivi
b. Dans la figure 1.A, le nombre de sujets à risque (No. at Risk) permet de déduire que 107 personnes sont décédées dans le groupe Empagliflozine durant les 6 premiers mois de l'étude (4 687 – 4 580 = 107)
c. Dans la figure 1.B, on observe que seuls 177 sujets sont encore suivis 4 ans après l'inclusion dans l'étude
d. Dans la figure 1.C, on observe que la médiane de survie à l'évènement « décès » dans le groupe Placebo est de 24 mois
e. Dans la figure 1.C, on dispose de suffisamment d'information pour affirmer que la différence de mortalité « toutes causes » entre le groupe Placebo et le groupe Empagliflozine est statistiquement significative

**11** **À propos de la figure 2 page 2124, indiquez quelle(s) affirmation(s) est (sont) correcte(s) ?**

a. Tous les résultats présentés sont issus d'analyses pré-spécifiées dans le protocole
b. Elle présente des analyses stratifiées
c. Les *p*-values ont été corrigées du fait de la multiplicité des tests réalisés présentés dans cette figure
d. L'effet observé pour le critère de jugement principal différait significativement entre les sujets âgés de plus ou moins 65 ans
e. L'effet global observé pour le critère de jugement principal était retrouvé quel que soit le niveau initial d'HbA1c

**12** **À propos du recueil des événements indésirables dans un essai clinique conduit en France, indiquez quelle(s) affirmation(s) est (sont) correcte(s) ?**

a. L'investigateur doit déclarer tout événement indésirable survenant au cours de l'essai au promoteur de l'étude
b. Un événement indésirable grave doit avoir entraîné une mise en jeu du pronostic vital du patient
c. Tout événement indésirable grave, même apparemment sans lien avec le médicament à l'étude, doit être déclaré
d. Toute suspicion d'effet indésirable grave inattendu doit être déclarée à l'ANSM
e. L'ANSM doit être informée immédiatement de tout évènement indésirable grave survenu lors d'un essai clinique



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**13** **À propos des conditions administratives afférentes aux essais cliniques et à la mise sur le marché d'un médicament en France, indiquez quelle est (quelles sont) la (les) réponse(s) exacte(s) ?**

a. L'Agence Nationale de Sécurité du Médicament (ANSM) est l'autorité compétente pour autoriser les essais cliniques
b. Dans le cadre des procédures d'Autorisations de Mise sur le Marché (AMM) nationales (demandes de commercialisations réservées uniquement à la France), l'autorité compétente pour délivrer l'AMM est la Haute Autorité de Santé (HAS)
c. En France, une fois délivrée, l'Autorisation de Mise sur le Marché n'est pas limitée dans le temps
d. Parmi les critères entrant en jeu dans l'inscription du médicament sur la liste des médicaments remboursables, on peut citer l'amélioration du service médical rendu
e. La Haute Autorité de santé fixe le prix du médicament

**14** **À propos des conclusions que vous pouvez tirer de cette étude, indiquez quelle est (quelles sont) la (les) réponse(s) exacte(s) ?**

a. L'empagliflozine est recommandée chez tous les sujets présentant un diabète de type 2
b. Malgré le nombre élevé d'essais cliniques antérieurs réalisés avec l'empagliflozine (*cf.* introduction), son profil d'évènements indésirables est encore mal connu
c. Cet article suggère que le choix de la posologie de l'empagliflozine doit être adapté aux valeurs de glycémie du patient
d. Cet article montre une relation dose-réponse entre l'empagliflozine et le critère de jugement principal
e. Les meilleurs résultats sur le critère de jugement composite sont attribuables à l'impact sur l'équilibre glycémique

**15** **À propos du résumé de cette étude, indiquez quelle est (quelles sont) la (les) réponse(s) exacte(s) ?**

a. L'objectif principal de cette étude est rapporté dans le résumé
b. Le critère principal de cette étude est défini dans le résumé
c. La marge de non-infériorité retenue dans cette étude est rapportée dans le résumé
d. La p-value de l'analyse principale de cette étude est rapportée dans le résumé
e. La conclusion du résumé apporte la réponse à l'objectif principal de cette étude

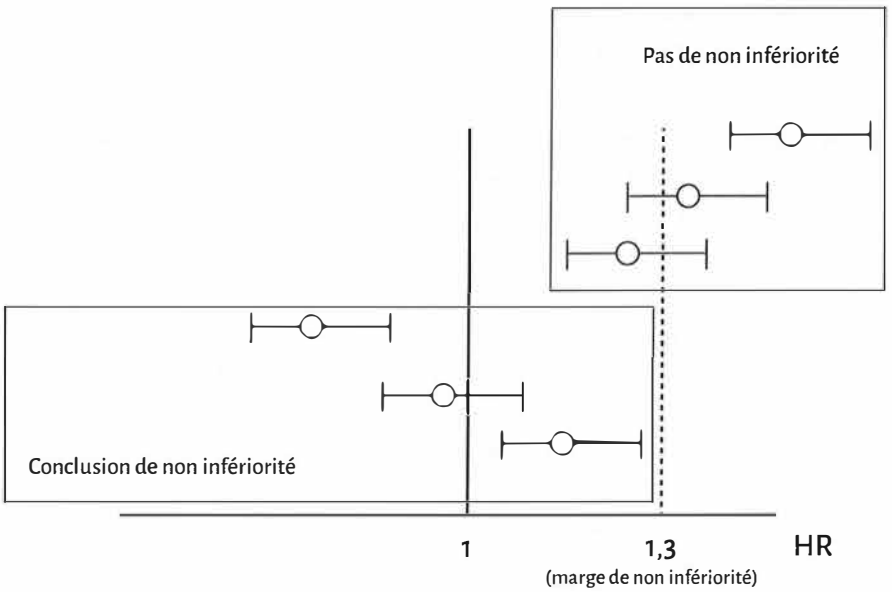


Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 1

**Réponse exacte : B**

*Commentaires* L'essai vise à tester la non-infériorité de l'empagliflozine par rapport au placebo sur le nombre de décès d'origine cardiovasculaire, d'IDM ou d'AVC avec une marge de non-infériorité de 1,3 sur le Hazard-ratio (HR). Autrement dit, l'empagliflozine sera déclarée non-inférieure au placebo si la borne supérieure de l'intervalle de confiance du Hazard-ratio est inférieure à 1,3. Par exemple, un HR de 1,10 avec un $IC_{95\,\%}$ = [1,02 ; 1,28] conduirait à déclarer l'empagliflozine comme non-inférieure au placebo. À l'inverse, un HR de 1,10 avec un $IC_{95\,\%}$ = [0,95 ; 1,36] conduira à conclure que l'empaglifozine n'est pas « non inférieure » au placebo puisque la valeur de 1,30 est incluse dans l'$IC_{95\,\%}$. Le schéma ci-dessous permet d'illustrer les différentes conclusions auxquelles aboutir selon les différentes étendues prises par l'$IC_{95\,\%}$ du HR :




## Question 2

**Réponses exactes : A – B – E**

*Commentaires*

A. Vrai. L'intervention consistant en l'administration, ou non, d'empagliflozine à 10 ou 25 mg/jour.

B. Vrai.

C. Faux. Un plan factoriel combine plusieurs interventions, permettant de répondre simultanément à plusieurs questions.

D. Faux. Un essai ouvert est un essai sans insu (insu est un synonyme d'aveugle).

E. Vrai.

## Question 3

**Réponse exacte : B**

*Commentaires*

A. Faux. il s'agit du nombre d'évènements (et non de sujets) correspondant au critère composite principal pour permettre de mettre en évidence la non-infériorité.

B. Vrai.

C. Faux. Il est de 0,0498.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

D. Faux. Il s'agit de bonnes pratiques relatives à l'éthique.

E. Faux. Rien n'indique que les analyses statistiques aient été réalisées en aveugle. La réalisation d'analyse « indépendante » ne signifie pas en aveugle.

### Question 4

**Réponse exacte : B**

*Commentaires*

A. Faux. Elle est comparée à un placebo.

B. Vrai (soit 10 soit 25 mg). Dès la 12e semaine, une adaptation posologique peut être réalisée pour les autres médicaments, mais pas pour l'empagliflozine.

C. Faux. Tous les sujets inclus reçoivent le placebo durant deux semaines (période de run-in), et la randomisation est ensuite réalisée.

D. Faux. L'objet d'une méta analyse est de combiner les données de plusieurs études répondant à la même question, dont les caractéristiques sont suffisamment convergentes pour pouvoir permettre l'agrégation des données.

E. Faux. On ne connaît pas l'effet de ce médicament sur les critères de jugement étudiés dans cet essai, on est donc en situation d'incertitude vis-à-vis de ces critères de jugement. De plus, cette étude a été réalisée pendant une période où les essais cités (réf. 13 à 20) étaient en cours de réalisation.

### Question 5

**Réponse exacte : E**

*Commentaires*

A. Faux. Le schéma initial prévoit 3 groupes, avec une randomisation 1 :1 :1, ce qui signifie qu'il y a autant de sujets dans chacun des 3 groupes (empagliflozine 10 mg, empagliflozine 25 mg, placebo), donc 2 fois plus recevant de l'empagliflozine qu'un placebo.

B. Faux. Ce n'est pas son objectif. La randomisation assure la comparabilité initiale des groupes et contribue ainsi à minimiser le phénomène de confusion.

C. Faux. La stratification permet d'assurer l'équilibre de la répartition des facteurs de stratification dans les 3 groupes initialement constitués. Elle garantit ici l'équilibre de certains facteurs pronostiques importants, dont il faut contrôler l'équilibre dans les groupes finalement analysés.

D. Faux. On peut contrôler son effectivité en comparant les distributions de plusieurs variables dans les différents groupes à l'inclusion dans l'étude (et non chez les groupes finalement analysés, moins nombreux (ITT modifiée), et pour lesquels il peut y avoir eu un phénomène d'attrition, c'est-à-dire la perte de vue au cours du suivi).

E. Vrai.

### Question 6

**Réponses exactes : A – B – C – D – E**

*Commentaires*

A. Vrai.

B. Vrai.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

C. Vrai. C'est souvent le cas et le risque d'interprétation excessive est fréquent. C'est ce qui est d'ailleurs observé dans l'analyse du critère de jugement secondaire avec l'angor instable, dont le risque était similaire sous empagliflozine ou placebo.

D. Vrai.

E. Vrai. Par exemple, dans cet essai, cette question se pose donc puisque l'exposition à l'empagliflozine augmente le risque d'AVC et diminue le risque d'infarctus du myocarde non fatal.

## Question 7

**Réponse exacte : B**

*Commentaires* Ces comités permettent l'étude et le classement des différents évènements (critères de jugement, évènements indésirables). Le fait qu'il en existe 2 distincts pour les événements cardiovasculaires et neurologiques ne contribue pas à augmenter le risque de biais de classement (également appelé biais de mesure) car ils analysent des événements distincts.

## Question 8

**Réponse exacte : E**

*Commentaires*

A. Faux. Il s'agit d'une analyse en ITT ('Intention To Treat') modifiée, puisqu'il fallait avoir reçu au moins une dose du médicament ou du placebo. Une analyse en intention de traiter consiste à analyser tous les patients dans leur groupe d'affectation, quel que soit le degré d'adhésion à l'intervention (c'est-à-dire dans le cas de cette étude, qu'ils aient réellement reçu ou non une dose du médicament ou du placebo). Une analyse en per protocol consiste à n'analyser que les patients qui ont effectivement suivi ce qui était prévu dans le protocole (en termes de traitement reçu, observance, durée).

B. Faux. Prendre en compte la mortalité comme un risque compétitif consiste à ne plus comptabiliser les patients décédés comme étant à risque de présenter l'évènement d'intérêt dès le jour de leur décès.

C. Faux. A ces 3 situations qui correspondent effectivement à une indication de censure, il faut ajouter un 4e cas : un décès de cause non cardiovasculaire (mortalité compétitive).

D. Faux. Par définition, aucune information postérieure à la censure n'est utilisée dans les analyses.

E. Vrai. Le risque alpha consenti ici est de 0,0498 du fait de cette correction pour la réalisation d'analyse intermédiaire.

## Question 9

**Réponse exacte : E**

*Commentaires*

A. Faux. Les chiffres donnés sont des taux pour 1 000 patients-années, c'est-à-dire que, pour 1 000 patients suivis un an dans le groupe placebo, on retrouvait en moyenne 43,9 fois le critère composite principal, contre 37,4 dans le groupe Empagliflozine.

B. Faux. Le risque était effectivement plus important dans le groupe Empagliflozine, puisque le HR est supérieur à 1, mais pas significativement plus important puisque l'intervalle de confiance à 95 % associé contient 1 ou, ce qui est équivalent, que la valeur critique « p » est supérieure à 0,05 (HR = 1,24, $IC_{95\ \%}$ = [0,92 ; 1,67], P = 0,16).

C. Faux. Il était significativement plus important dans le groupe Placebo (HR = 0,65, $IC_{95\ \%}$ = [0,50 ; 0,85], P = 0,002).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

D. Faux. L'ischémie myocardique silencieuse (IMS) doit être spécifiquement investiguée, ce qui suppose une contrainte supplémentaire pour les patients, on s'attend donc à ce que cette information ne soit disponible que pour une sous-partie de l'effectif. Ici, 38 patients présentent bien une IMS, mais la note en bas du tableau précise que l'IMS a été recherchée chez seulement 2 378 patients du groupe empagliflozine. On retrouve bien 38/2 378 ≈ 1,6 %.

E. Vrai. Un HR (Hazard-Ratio, traduisible ici par rapport de risque) à 0,68 implique une réduction du risque de 32 % (1-0,68 = 0,32).

## Question 10

**Réponses exactes : A – E**

*Commentaires*

A. Vrai.

B. Faux. Le nombre de sujets à risque indique le nombre de sujets encore analysés aux différents temps, mais un sujet peut ne plus être analysé pour d'autre raisons, par exemple ici en particulier un infarctus du myocarde non fatal ou un accident cardiovasculaire (AVC) non fatal. On peut donc seulement dire que 107 individus du groupe Empagliflozine ont rencontré l'évènement ou ont été censurés pour une autre raison dans cet intervalle de temps.

C. Faux. C'est le nombre de sujets encore suivis dans le groupe placebo.

D. Faux. La médiane de survie correspond au temps où la fonction de survie vaut 50 %. Or ici il s'agit de courbes de mortalité calculées par l'estimation de Kaplan-Meier (1 – la fonction de survie), et elle ne dépasse pas 15 % à la fin des 4 ans. La survie atteint donc environ 85 % après 4 ans. La médiane de survie à l'évènement « décès » n'est pas connue, on sait simplement qu'elle est postérieure au suivi, et ne peut être déduite du graphique ni des données présentées.

E. Vrai. Le Hazard ratio (HR) est donné avec son intervalle de confiance, qui ne contient pas 1. À noter toutefois que le modèle de Cox qui a permis son calcul est ajusté sur une série de facteurs nécessaires à une bonne compréhension de ces chiffres, mais non donnés dans la figure ou sa légende.

## Question 11

**Réponses exactes : B – D**

*Commentaires*

A. Faux. Les analyses en sous-groupes pour le critère de jugement principal étaient spécifiées dans le protocole mais pas les analyses sur le décès d'origine cardiovasculaire. Ces dernières sont des analyses post-hoc (c'est-à-dire qu'elles n'étaient pas prévues dans le protocole de l'étude).

B. Vrai. Il s'agit d'analyses stratifiées où l'effet est présenté séparément dans chacune des strates (par ex. dans la strate des < 65 ans et dans la strate des 65 ans ou plus).

C. Faux. Cela est précisé dans la légende de la figure (« with no adjustment for multiple testing »).

D. Vrai. Le test d'interaction est significatif (p = 0,01).

E. Faux. L'effet semblait être différent en fonction du niveau initial d'HbA1c (test d'interaction significatif).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 12

**Réponses exactes : A – C – D**

*Commentaires*

A. Vrai (dans le cahier d'observation de l'étude).

B. Faux. Un événement indésirable est qualifié de grave s'il

- entraîne un décès ;
- met en jeu le pronostic vital ;
- nécessite une hospitalisation ou la prolongation de l'hospitalisation ;
- provoque une incapacité ou un handicap important ou durable ;
- se traduit par une anomalie ou une malformation congénitale.

C. (*Cf.* Article R-1123-46 du code de la santé publique).

D. Vrai. L'investigateur notifie au promoteur sans délai à compter du jour où il en a connaissance tous les événements indésirables graves ainsi que tous les effets indésirables et les incidents graves survenus au cours d'une recherche mentionnée au 1° de l'article L. 1121-1, à l'exception de ceux qui sont recensés dans le protocole ou dans la brochure pour l'investigateur comme ne nécessitant pas une notification sans délai (Article R-1123-49 du code de la santé publique). L'investigateur doit donc notifier au promoteur tous les événements indésirables graves.

Un événement indésirable correspond à toute manifestation nocive survenant chez une personne qui se prête à une recherche impliquant la personne humaine, que cette manifestation soit liée ou non à la recherche ou au produit sur lequel porte cette recherche (Article R-1123-46 du code de la santé publique) Un événement indésirable n'est pas forcément en lien avec le produit à l'étude.

E. Vrai. Déclaration à faire sans délai pour les effets indésirables graves inattendus ayant entraîné la mort ou mis la vie en danger, dans un délai de 15 jours pour les autres effets indésirables graves inattendus. Un effet indésirable inattendu se définit comme « tout effet indésirable du produit dont la nature, la sévérité, la fréquence ou l'évolution ne concordent pas avec les informations de référence sur la sécurité mentionnées dans le résumé des caractéristiques du produit ou dans la brochure pour l'investigateur lorsque le produit n'est pas autorisé ».

F. Faux. Le promoteur doit informer l'ANSM sans délai ou sous un délai de 15 jours de tout effet indésirable grave inattendu. Concernant les événements indésirables graves non inattendus, l'ANSM est informée annuellement dans le rapport de sécurité de l'étude.

## Question 13

**Réponse exacte : A**

*Commentaires*

A. Vrai.

B. Faux. C'est l'Agence Nationale de Sécurité du Médicament (ANSM) qui délivre l'AMM en cas de procédure nationale (la demande d'AMM ne concerne que la France et n'est pas faite à l'échelle européenne). Ces procédures sont de plus en plus rares au profit des procédures communautaires concernant l'ensemble de l'Europe, l'autorité compétente étant alors l'Agence Européenne du Médicament (EMA).

C. Faux. Elle est réévaluée systématiquement à 5 ans puis donnée à vie, sous réserve d'une réévaluation possible au moindre besoin (nouvelles données de sécurité).

D. Faux. L'inscription du médicament se fait sur la base du service médical rendu (SMR), son prix sur la base de l'amélioration du service médical rendu (ASMR).

E. Faux. C'est le rôle du Comité Économique des Produits de Santé (CEPS).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 14

**Réponse exacte : B**

*Commentaires*

A. Faux. Il faut bien regarder les critères de sélection pour proposer une population cible, ici entre autres les sujets de plus de 18 ans, présentant un diabète de type 2, avec un IMC inférieur à 45 kg/m$^2$, une HbA1c entre certaines valeurs cibles, et une insuffisance rénale chronique sévère...

B. Vrai. Seuls les événements indésirables fréquents et à court terme sont détectables durant les essais cliniques.

C. Faux. Cet aspect n'est pas étudié dans l'article. En particulier, la dose d'empagliflozine n'était pas adaptée au patient puisqu'elle était attribuée en aveugle.

D. Faux. Au contraire, les auteurs insistent sur le fait que les résultats semblent similaires avec les deux doses utilisées, en particulier pour les rapports de risque pour le critère de jugement principal, avec pour 10 mg un HR à 0,85 ($IC_{95\,\%}$, [0,72 ; 1,01]) et pour 25 mg un HR à 0,86 ($IC_{95\,\%}$, [0,73 ; 1,02]) par rapport au placebo.

E. Faux. On ne peut que spéculer sur les multiples causes qui ont pu conduire à cette diminution de la mortalité cardiovasculaire, et l'équilibre du diabète est une piste parmi d'autres (amélioration de la fonction cardiaque, diminution de la rigidité artérielle, impact sur le système rénine-angiotensine aldostérone...).

## Question 15

**Réponse exacte : B**

*Commentaires*

A. Faux. Ainsi, ce résumé ne correspond pas aux recommandations de bonnes pratiques de rédaction d'un article relatif à un essai clinique, tel qu'attendu dans la grille CONSORT.

B. Vrai.

C. Faux.

D. Faux.

E. Faux.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 8 Essai en cross-over

Damien Jolly, Moustapha Dramé

| | |
|---|---|
| **Titre de l'article** | **Glycomacropeptide for nutritional management of phenylketonuria : a randomized, controlled, crossover trial** |
| **Auteurs** | Ney DM, Stroup BM, Clayton MK, Murali SG, Rice GM, Rohr F, Levy HL |
| **Référence** | Am J Clin Nutr. 2016 Aug ; 104(2):334-45 |
| **Lien vers l'article** | http://ajcn.nutrition.org/content/104/2/334.long |

**1** **Les glycomacropeptides sont**

a. À la base de nombreux compléments alimentaires utilisables dans la phénylcétonurie
b. Des composés à faible teneur en phénylalanine
c. Des composés riches en protéines
d. Des composés isolés à partir du petit lait
e. Des composés possédant des propriétés probiotiques

**2** **Les études préliminaires animales concernant le glycomacropeptide, qui comparaient notamment les apports en acides aminés, ont montré**

a. Qu'il favorisait la croissance
b. Qu'il réduisait les concentrations de phénylalanine dans le plasma et dans le cerveau
c. Qu'il améliorait le statut osseux
d. Qu'il permettait d'atténuer le stress métabolique
e. Qu'il améliorait la fonction rénale

**3** **Chez l'Homme, les médicaliments riches en glycomacropeptide utilisés dans une étude préliminaire ont montré**

a. Qu'ils pouvaient être utilisés en toute sécurité dans la phénylcétonurie
b. Qu'ils étaient bien acceptés par les patients
c. Qu'ils donnaient une impression de plus grande satiété
d. Qu'ils permettaient d'apporter plus de protéines
e. Qu'ils permettaient d'optimiser les apports en protéines en limitant les apports en médicaliments riches en acides aminés

**4** **Concernant la typologie de l'étude réalisée, on peut la qualifier**

a. D'essai randomisé
b. D'essai comparatif multicentrique
c. D'essai où chaque sujet est pris comme son propre témoin
d. D'essai d'efficacité
e. D'essai de tolérance



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**5** **Le choix du design en cross-over a été fait**

a. Car il s'agit d'un mode de comparaison « élégant »
b. Car les sujets atteints de phénylcétonurie sont différents dans leurs formes cliniques et dans leurs gravités
c. Car les sujets atteints de phénylcétonurie ont tous des régimes différents spécifiques, adaptés à leur mode de vie
d. Car le nombre de sujets porteurs de phénylcétonurie inclus correspond à la moitié du nombre de mesure du critère de jugement principal
e. Car le nombre de sujets porteurs de phénylcétonurie dans ce type d'essai est plus faible que dans un essai en deux groupes parallèles

**6** **Les objectifs de l'étude**

a. Étaient de comparer deux modes de supplémentation en acides aminés et en protéines rendue nécessaire par le régime d'exclusion de phénylcétonurie permettant de limiter les apports en phénylalanine
b. Étaient de tester un nouveau type d'apport en acides aminés, en protéines et en micro-nutriments pour éviter certaines atteintes neurocognitives, des fragilités osseuses et des atteintes rénales connues chez les utilisateurs à long terme de médicaliments riches en acides aminés
c. Étaient de comparer deux types de régime pauvre en phénylalanine
d. Étaient de comparer deux types de supplémentation en acides aminés et en protéines permettant les mêmes apports alimentaires mesurés en équivalents/protéine
e. Étaient de réaliser un essai en conditions réelles de vie des patients, en ambulatoire

**7** **Parmi les critères mesurés dans cette étude, on trouvait**

a. Une étude du profil en acides aminés du plasma
b. Un recueil des apports alimentaires
c. Des tests neuropsychologiques pour évaluer les fonctions exécutives
d. Des mesures du coefficient intellectuel
e. Des mesures de la dépression grâce à une échelle spécifique

**8** **Globalement l'essai**

a. Se déroulait sur 22 semaines
b. Comportait une période de wash out de 3 semaines
c. Comportait pour chaque patient deux périodes d'une semaine d'éducation
d. Comportait deux périodes de traitement de 4 semaines
e. Comportait 4 mesures de l'acceptabilité du traitement reçu, au cours du suivi



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**9** **Du point de vue de l'analyse dont les résultats sont présentés, différents effets ont été testés**

a. Un effet groupe de tirage au sort (ou séquence des régimes thérapeutiques) mesurant l'interaction entre l'effet traitement et l'effet période de traitement (comparant l'effet du traitement reçu en premier à celui du traitement reçu en second)
b. Un effet période de traitement (comparant l'effet du traitement reçu en premier à celui du traitement reçu en second, interaction entre le groupe de tirage au sort (ou séquence des régimes thérapeutiques) et l'effet du traitement
c. Un effet traitement permettant de voir s'il existe des différences liées au type de médicaliments de supplémentation en protéines et en acides aminés
d. Un effet carry-over ou effet testant l'influence du premier traitement sur le traitement administré en second
e. Un effet direct du temps, comparant les effets des traitements mesurés sur les différences entre J21 et J0 pour chaque médicaliment testé

**10** **Parmi les 173 sujets éligibles à l'essai, après analyse des dossiers des hôpitaux**

a. 17 sujets étaient perdus de vue
b. 17 sujets n'étaient pas intéressés pour participer à l'essai
c. 8 sujets n'aimaient pas manger des aliments médicaux contenant des glycomacropeptides
d. 32 sujets ont été inclus
e. 2 sujets n'ont pas été capables de finir l'étude

**11** **Concernant le critère de jugement principal : les variations des concentrations plasmatiques de phénylalanine**

a. Les comparaisons ont été faites sur les différences entre les concentrations à J22 et les concentrations initiales à J1 pour chacune des périodes de traitement
b. Les comparaisons ont été faites en ajustant sur les niveaux initiaux de phénylalanine mesurés à J1 pour chacun des sujets et pour chaque période de traitement
c. Les comparaisons ont été faites en ajustant sur les niveaux d'apports alimentaires moyens en phénylalanine de chacun des sujets
d. Il existait un effet traitement : une différence significative entre les deux groupes de traitement (p = 0,0008)
e. Il n'existait pas d'effet séquence thérapeutique ni d'interaction séquence-traitement



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**12** **En pratique, les effets du type de médicaliment de supplémentation en acides aminés et protéines ont montré**

a. Une différence moyenne significative de 147 µmol/L de la phénylalanine plasmatique entre les deux groupes médicaliments riches en acides aminés et médicaliments riches en glycomacropeptides, défavorable au groupe glycomacropeptides
b. Une augmentation moyenne non significative de 62 µmol/L de la phénylalanine plasmatique entre la phénylalanine mesurée à J22 et celle mesurée à J1 lors de la supplémentation avec médicaliments riches en glycomacropeptides
c. Une diminution moyenne significative de 85 µmol/L de la phénylalanine plasmatique entre la phénylalanine mesurée à J22 et celle mesurée à J1 lors de la supplémentation par des médicaliments riches en acides aminés
d. Une augmentation significative des apports en phénylalanine alimentaire dans le groupe médicaliments riches en glycomacropeptides évaluée à environ 184 mg par jour en moyenne
e. Une augmentation moyenne non significative de 99 mg par jour des apports en phénylalanine dans le groupe médicaliments riches en acides aminés

**13** **Le ratio phénylalanine sur tyrosine est souvent utilisé cliniquement comme une mesure du contrôle de la phénylalaninémie dans la phénylcétonurie. Dans cette étude, on retrouve**

a. Un effet traitement significatif à J22 après ajustement sur le ratio basal mesuré à J1 avec une tendance à l'augmentation dans le groupe médicaliments riches en glycomacropeptides
b. Une tendance à l'augmentation du ratio en moyenne de 2,42 dans le groupe médicaliments riches en glycomacropeptides
c. Une diminution significative du ratio en moyenne de 3,65 pour le groupe médicaliments riches en acides aminés
d. Un effet traitement significatif sur l'analyse brute
e. Des effets en accord avec les variations observées dans le dosage plasmatique de la phénylalanine et de la tyrosine

**14** **Du point de vue de la tolérance des suppléments alimentaires testés**

a. Le comité de sécurité n'a pas mis en évidence d'effet secondaire provoqué par les médicaliments testés
b. Après trois semaines de traitement les effets sur la fonction rénale étaient similaires dans les deux groupes
c. Après trois semaines de traitement les effets sur les électrolytes étaient similaires dans les deux groupes
d. Après trois semaines de traitement les niveaux d'albuminémie étaient plus élevés avec une augmentation significative de l'albumine dans le groupe traité par glycomacropeptides en faveur d'un meilleur statut nutritionnel
e. Après trois semaines de traitement les niveaux d'enzymes mesurant les fonctions hépatiques étaient similaires dans les deux groupes



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**15** **En termes de conclusion**

a. Les auteurs affirment que les médicaliments riches en glycomacropeptides sont utilisables dans le traitement de la phénylcétonurie en suggérant parallèlement de réduire la teneur en phénylalanine de ces produits.
b. Les auteurs affirment que l'augmentation non significative entre le niveau en fin de traitement et le niveau basal dans le groupe glycomacropeptide est un argument pour dire que le glycomacropeptide est utilisable dans le traitement de la phénylcétonurie.
c. Les auteurs utilisent comme argument complémentaire en faveur du glycomacropeptide le fait qu'il apporte plus de protéines, qu'il est mieux toléré et qu'il permet un meilleur statut nutritionnel
d. Les affirmations des auteurs sont abusives car le nouveau traitement montre une augmentation significative des taux de phénylalaninémie potentiellementtoxiques après trois semaines de traitement, et pouvant donc entraîner à long terme des complications déjà bien connues
e. La conclusion des auteurs en faveur du traitement par médicaliment riche en glycomacropeptides est d'autant plus abusive qu'elle se base sur une observation de trois semaines alors que le traitement proposé est un traitement à vie, et que l'on ne peut absolument pas extrapoler les effets secondaires à long terme



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 1**

**Réponses exactes : A – B – C – D**

*Commentaires* La réponse E est fausse : les glycomacropeptides possèdent des propriétés prébiotiques et non probiotiques. Tout est rapporté dans la partie contexte de l'article.

**Question 2**

**Réponses exactes : A – B – C – D**

*Commentaires* La réponse E est fausse car non citée dans le texte.

**Question 3**

**Réponses exactes : A – B – C – D – E**

**Question 4**

**Réponses exactes : A – B – C – D – E**

*Commentaires* L'ensemble de ces réponses sont citées nommément dans le texte de l'article. L'essai est bien multicentrique puisqu'il associait deux centres : le *Waisman center* et le *Boston children hospital*, et comparatif car on compare 2 « traitements », 2 groupes de sujets, et 2 mesures du critère de jugement chez un même sujet.

**Question 5**

**Réponses exactes : B – C – D – E**

*Commentaires* Réponse A : Ce n'est pas une raison suffisante même si c'est vrai. Les réponses B et C sont vraies car dans ces cas le design en cross-over améliore la comparabilité des groupes. La réponse D est vraie car en plus la maladie est une maladie rare, et donc la diminution du nombre de sujets par rapport à un design classique favorisera l'inclusion dans la phase finale. La réponse E est la vraie raison de l'utilisation de ce design particulier en cross-over.

**Question 6**

**Réponses exactes : A – B – D – E**

*Commentaires* La réponse C est fausse puisque les comparaisons qui sont faites concernent les suppléments alimentaires et non pas le régime pauvre en phénylalanine qui reste spécifique à chaque sujet, et n'est pas modifié.

**Question 7**

**Réponses exactes : A – B – C**

*Commentaires* A, B, C : Ces données sont citées dans le paragraphe *Study design and protocol* de l'article. Réponse D : La mesure du coefficient intellectuel n'était faite que si elle manquait dans le dossier clinique des patients. Réponse E : L'affirmation est fausse car c'est l'anxiété qui était mesurée.

**Question 8**

**Réponses exactes : B – C – E**

*Commentaires* La réponse A est fausse, c'est 11 semaines. La réponse D est fausse car les périodes de traitement étaient des périodes de 3 semaines.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 9**

**Réponses exactes : A – B – C – E**

*Commentaires* Réponse D : Cet effet est évité *a priori* par la période de wash-out qui est relativement longue.

**Question 10**

**Réponses exactes : A – B – D – E**

*Commentaires* Tous ces résultats sont donnés sur le diagramme de flux (*cf.* figure 2).

**Question 11**

**Réponses exactes : A – B – C – D – E**

*Commentaires* Les réponses A, B, C sont vraies : cela est expliqué dans le 3e paragraphe de l'analyse statistique. Les réponses D et E sont vraies : Cela est expliqué en haut de la 2e colonne de la page 7 de l'article.

**Question 12**

**Réponses exactes : A – B – C – D – E**

*Commentaires* Toutes ces données peuvent se retrouver dans la figure 4.

**Question 13**

**Réponses exactes : A – B – C – E**

*Commentaires* Réponse D : Cette affirmation est fausse. Il n'y a pas d'effet traitement sur le critère de jugement secondaire en analyse non ajustée. Les résultats de l'analyse brute ne sont pas donnés.

**Question 14**

**Réponses exactes : A – B – C – D – E**

*Commentaires* La réponse E est expliquée page 6 2e colonne du 2e paragraphe.

**Question 15**

**Réponses exactes : A – B – C – D – E**

*Commentaires* La réponse A est fausse : ces médicaliments à base de glycomacropeptides augmentent les niveaux sanguins de phénylalanine. Ce n'est donc pas bon pour le cerveau du patient !

La réponse B est fausse : l'analyse de l'essai en cross-over ne permet pas ce type d'analyse.

Les auteurs sont bien conscients des problèmes, puisqu'ils suggèrent eux-mêmes de réduire la concentration de phénylalanine des médicaliments à base de glycomacropeptides.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 9 Essai en cross-over

Damien Jolly

| | |
|---|---|
| **Titre de l'article** | **Spironolactone *versus* placebo, bisoprolol, and doxazosin to determine the optimal treatment for drug-resistant hypertension (PATHWAY-2) : a randomised, double-blind, cross over trial** |
| **Auteurs** | Williams B, MacDonald TM, Morant S, Webb DJ, Sever P, McInnes G, Ford I, Cruickshank JK, Caulfield MJ, Salsbury J, Mackenzie I, Padmanabhan S, Brown MJ, British Hypertension Society's PATHWAY Studies Group |
| **Référence** | Lancet. 2015 Nov 21 ; 386(10008):2059-68 |
| **Lien vers le site** | **https://doi.org/10.1016/S0140-6736(15)00257-3** |

**1** **De quel type d'étude s'agit-il ?**

a. Une étude thérapeutique contrôlée
b. Un essai comparant trois groupes de traitements antihypertenseurs
c. Un essai où chaque sujet est pris pour son propre témoin
d. Un essai comportant 24 groupes de séquences thérapeutiques différentes
e. Un essai qui devrait permettre de connaître les efficacités anti hypertensives additionnelles respectives de chacun des trois traitements testés par rapport au placébo et les uns par rapport aux autres

**2** **On peut dire que cette étude**

a. A permis une augmentation de doses au double de la dose initiale, six semaines après l'initiation de chaque principe actif
b. A permis de traiter les sujets pendant un an
c. A bénéficié d'une période de standardisation des traitements des patients, avec un mois initial de traitement, identique pour tous les patients
d. A permis de rajouter un 4$^{e}$ traitement chez les patients pendant 36 semaines
e. A amené 6 sous-groupes de patients, avec une HTA résistante, à ne recevoir aucun traitement actif, pendant 4 mois

**3** **Dans cette étude, l'objectif principal était**

a. De déterminer si la spironolactone est globalement le meilleur médicament à rajouter chez les patients porteur d'une HTA résistante
b. De tester si l'HTA résistante était bien liée à une surcharge en sodium
c. De montrer que la spironolactone était plus efficace qu'un placebo
d. De montrer que la spironolactone était plus efficace que la Doxazosine
e. De montrer que la spironolactone était plus efficace que le bisoprolol



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **Dans cette étude, la population étudiée était**

a. Constituée de patients hypertendus avec une HTA résistante à un traitement bien mené depuis trois mois selon les recommandations internationales
b. Constituée de patients issus de 12 centres de soins secondaires
c. Constituée de 335 patients dont seulement 230 ont reçus les quatre traitements prévus a été répartie par tirage au sort en groupe d'environ 14 patients constituée de 436 patients présélectionnés, dont 88 ont été exclus sans en connaître les raisons réelles
d. A été répartie par tirage au sort en groupe d'environ 14 patients
e. Constituée de 436 patients présélectionnés, dont 88 ont été exclus sans en connaître les raisons réelles

**5** **Dans cette étude, le critère du jugement**

a. N'est pas un critère classiquement utilisé
b. Est un critère composite
c. Est une estimation de la Pression artérielle systolique (PAS) moyenne calculée à partir de la PAS auto-mesurée à domicile, à partir de trois mesures interactives, chaque matin et chaque soir, pendant 4 jours consécutifs est une estimation de la PAS moyenne : PAS auto-mesurée à domicile, disponibles sur les trois derniers jours avant la visite
d. Est une estimation de la PAS moyenne auto-mesurée à domicile, à partir d'au moins six mesures disponibles sur les trois derniers jours avant la visite
e. Est une estimation de la PAS moyenne auto-mesurée à domicile, à partir d'au moins six mesures disponibles sur les trois derniers jours avant la visite

**6** **Dans cette étude, du point de vue des statistiques**

a. Le nombre de sujets nécessaire a été calculé à 314
b. 294 patients ont été inclus
c. Le risque alpha a été fixé à 0,003 pour le calcul du nombre de sujet nécessaires
d. Le risque bêta a été fixé à 0,10
e. Le risque alpha final a été inutilement fixé à 0,01 puisque l'analyse *a priori* hiérarchique protégeait de l'inflation du risque alpha

**7** **Dans cette étude, les comparaisons de PAS moyennes au sein des groupes traités**

a. Ont été faites par des tests t de Student
b. Ont été faites grâce à un modèle mixte pour mesures répétées
c. Ont été ajustées sur la PAS initiale de chacun des patients
d. Ont été ajustées sur les caractéristiques de chaque patient (âge, sexe, taille, poids, habitudes tabagiques)
e. Ont été faites sur des moyennes estimées par la méthode des moindres carrés



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**8** **Dans les tableaux 2 et/ou 3 du chapitre « résultats », les auteurs**

a. Présentent l'essentiel des résultats, en termes d'effets sur le critère de jugement principal
b. Montrent les résultats des analyses hiérarchisées successives prévues
c. Donnent des estimations des ampleurs des effets globaux de chacun des médicaments utilisés quelle que soit la dose utilisée
d. Donnent des estimations des ampleurs des effets globaux de chacun des médicaments utilisés à pleine dose
e. Considèrent les résultats du tableau 3 comme une analyse de sensibilité des résultats présentés dans le tableau 2

**9** **Les auteurs avaient décidé d'utiliser comme critère de jugement secondaire, l'atteinte d'une PAS cible sous traitement**

a. La PAS cible était définie comme une PAS moyenne calculée inférieure à 130 mm Hg les sujets hypertendus résistants aux traitements
b. La PAS cible était définie comme une PAS moyenne calculée inférieure à 135 mm Hg les sujets hypertendus résistants aux traitements
c. La PAS cible était définie comme une PAS moyenne calculée inférieure à 125 mm Hg chez les sujets hypertendus résistants aux traitements mais diabétiques
d. Les pourcentages de sujets obtenant un contrôle de PAS au sein de chacun des groupes ont été comparés par ses modèles de régression logistique
e. Ces résultats en termes d'atteinte d'une PAS cible, confirment les résultats observés sur les PAS moyenne avec la spironolactone

**10** **À propos des résultats figurant dans le tableau 2**

a. Sous placebo, la pression artérielle moyenne prédite par modèle mixte a 90 % de chances de se situer entre 142,6 et 114,6 mmHg
b. La pression artérielle moyenne prédite est réduite en médiane de 12,8 mmHg sous spironolactone
c. La spironolactone augmente en moyenne de 8,70 mmHg la pression artérielle prédite par rapport au placebo
d. La spironolactone a 97,5 % de chance de réduire d'au moins 3,46 mmHg la pression artérielle par rapport au bisoprolol
e. La spironolactone a 25 % de chances de réduire de plus de 5,04 mmHg la pression artérielle par rapport à la daxazosin

**11** **Les auteurs avaient décidé d'utiliser comme critère de jugement secondaire, la chute de PAS moyenne en fonction des niveaux de rénine des patients**

a. Les niveaux de rénine ne sont pas liés à l'ampleur de la chute de PAS sous Bisoprolol
b. Les niveaux de rénine ne sont pas liés à l'ampleur de la chute de PAS sous Doxazosine
c. Les niveaux de rénine sont statistiquement liés à l'ampleur de la chute de PAS sous Spironolactone
d. Les niveaux de rénine n'expliquent quasiment pas l'ampleur de la chute de PAS sous Spironolactone, expliquant seulement moins de 4 % de la variance expérimentale de cette chute
e. La figure 3 tend à laisser supposer qu'au-dessus d'un certain niveau de rénine plasmatique, le bisoprolol serait plus efficace que la spironolactone



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**12** **Les auteurs avaient décidé d'utiliser comme critère secondaire la fréquence de la survenue d'effets secondaires sous traitement**

a. Les principes actifs ont globalement été bien tolérés
b. Les proportions d'effets secondaires graves sont quasi identiques (non différents) entre les trois groupes traités et le groupe sous placebo
c. Pris globalement, quel que soit le type d'effet secondaire, les évènements indésirables graves sont significativement plus fréquents dans l'un des groupes traités comparé au groupe placebo
d. On peut affirmer qu'il existe une différence significative en termes de pourcentages d'effets secondaires, quel que soit le type d'effet secondaire, entre les groupes Doxazosine et placebo
e. On peut affirmer qu'il existe une différence significative en termes de pourcentages d'effets secondaires, quel que soit le type d'effet secondaire, entre les groupes spironolactone et placebo

**13** **En termes de validité, cette étude**

a. Est une des premières études faites sur une population de sujets avec une HTA résistante, bien caractérisée
b. Possède un schéma expérimental original ayant permis des comparaisons simultanées et multiples entre les trois médicaments antihypertenseurs
c. Est la deuxième étude, utilisant des moyennes de mesure de pression artérielle à domicile,
d. Dont la durée des traitements pendant seulement trois mois reste une limite en termes de généralisation des résultats
e. Permet de proposer la spironolactone comme traitement ajouté à une trithérapie bien menée chez ces patients avec HTA résistante

**14** **La réalisation de ce cross over**

a. A permis globalement de gagner sur le nombre de sujet à inclure
b. A donc facilité la réalisation de l'étude
c. A sans doute permis de gagner du temps sur l'obtention de la réponse à la question énoncée en hypothèse
d. N'était possible que du fait de la chronicité et de la non réversibilité de cette situation d'HTA résistante, traité par déjà trois traitements standardisés, et résistant à ceux–ci
e. N'a d'intérêt que si l'on accepte que la variabilité de pression artérielle, entres les sujets porteurs d'une HTA résistante, sélectionnés sur des critères précis, et traités de la même manière en trithérapie efficace est plus petite que la variabilité de la pression artérielle chez un même sujet en fonction des moments de la journée

**15** **Les limites de cette étude sont**

a. L'exclusion des patients avec un débit de filtration glomérulaire inférieur à 45 ml/min
b. L'absence de période de wash out
c. L'inclusion essentiellement de patients caucasiens
d. L'absence d'évaluation de la morbidité et de la mortalité
e. Le design de l'étude en cross over



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 1

**Réponses exactes : A – B – C – D – E**

*Commentaires*

A : L'étude est contrôlée car le groupe placebo sert bien de référence pour la comparaison des effets des traitements. Les autres traitements servent aussi de comparateurs.

B et C : Chaque sujet reçoit en théorie successivement et dans un ordre aléatoire un placebo et trois traitements antihypertenseurs issus de trois classes thérapeutiques différentes, donnés en simple dose puis en double dose pendant deux phases successives de 6 semaines.

D : C'est un cross over original avec quatre groupes de traitements pris successivement par chaque patient en complément de sa trithérapie anti-hypertensive habituelle ; ainsi 24 combinaisons de séquences thérapeutiques étaient possibles.

E : La stratégie hiérarchique d'analyse va donc permettre de situer chaque diminution de pression artérielle systolique moyenne, pour chaque médicament comparé au placebo, puis en comparaison des effets de la spironolactone à l'effet moyen des deux autres drogues, puis éventuellement à chacun des médicaments isolément.

### Question 2

**Réponses exactes : A – B – D**

*Commentaires*

A : *cf.* texte

B : 4 semaines de run-in sous placebo, puis dans un ordre aléatoire, 6 semaines à demi-dose, puis 6 semaines à pleine dose, cela répété quatre fois soit 48 semaines, ce qui amène à 52 semaines... Même sous placebo, on peut considérer que ces patients porteurs d'une HTA résistante, sont « traités », puisqu'il s'agit de l'ajout d'un quatrième médicament. Les patients attribués à la séquence commençant par le placebo après le run-in, fait sous placebo, ont donc reçu 16 semaines de placebo, mais en conservant leur traitement associant trois médicaments habituels, un inhibiteur de l'enzyme de conversion de l'angiotensine ou un bloqueur des récepteurs de l'angiotensine II, un inhibiteur calcique, un thiazidique ou une thiazidique like...

C Pendant le run-in, les patients conservent leurs traitements habituels mais ces traitements ne sont pas identiques.

D : Trois traitements efficaces pendant 12 semaines chacun soit 36 semaines.

E : Les patients conservaient leur trithérapie habituelle.

### Question 3

**Réponse exacte : A**

*Commentaires*

A : C'est l'objectif énoncé tel quel dans le texte. (paragraphe 1 page 2061)

B à E : Les autres affirmations sont fausses car ce ne sont pas des objectifs principaux, même si les trois dernières affirmations se révèlent justes en termes de résultats...

### Question 4

**Réponses exactes : A – C – D – E**

*Commentaires*

A : C'est énoncé tel quel dans le texte (page 2063 premier paragraphe des résultats) et dans le flow chart.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

B : 12 centres de soins secondaires et 2 centres de soins primaires ont participés

C et D : 335 patients, répartis en 24 groupes par tirage au sort, donne un effectif moyen de 13,9 ou 14 patients par groupe.

E : Plus de 23 % d'exclusions sans en connaître les raisons détaillées pose un problème de fond sur la validité de l'étude.

### Question 5

**Réponses exactes : A – D – E**

*Commentaires*

A : Oui l'auto mesure de la PAS à domicile est un mode original de mesure de la PAS en vraie vie.

B : Ce n'est pas un critère composé de plusieurs sous critères élémentaires.

C : Le critère de jugement est estimé sur seulement 3 jours : i. e. J2 J3 J4 (c. texte page 2061, 2e colonne, 2e paragraphe).

D et E : Ce qui énoncé dans les deux dernières affirmations, est juste.

### Question 6

**Réponses exactes : C – D – E**

*Commentaires* Les chiffres des propositions A et B ont été inversés. Ces éléments sont donnés par une lecture et une compréhension du premier paragraphe du chapitre analyse statistique. (page 2062).

### Question 7

**Réponses exactes : B – C – D – E**

*Commentaires*

A : Non, le test t a juste été utilisé dans le calcul du nombre de sujets nécessaire.

B : Oui, c'est plutôt un modèle proche de la régression linéaire généralisée, prenant en comptes les variables de PAS mesurées au cours du temps comme des variables à expliquer en fonction du temps et donc des différentes périodes de traitement. L'auto mesure de la PAS à domicile est un mode original de mesure de la PAS en vraie vie.

Dans les variables explicatives, la PAS de base et les caractéristiques de base de chaque patient ont été forcées (Item D) pour ajuster les moyennes estimées par la méthode des moindres carrés (Item E).

### Question 8

**Réponses exactes : A – B – C – D – E**

*Commentaires* Toutes les réponses sont justes, *cf.* texte dans méthodes et résultats.

### Question 9

**Réponses exactes : B – D – E**

*Commentaires*

Les items A et C sont des leurres.

B : il s'agit bien de la valeur cible indiquée, *cf.* texte.

E : vrai, *cf.* texte.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 10

**Réponses exactes : C – D**

*Commentaires*

A : l'intervalle de confiance est à 95 % et non à 90 %.

B : il s'agit d'une moyenne et non d'une médiane.

C : il y a visiblement une erreur de signe pour cette valeur ; l'intervalle de confiance ne comprend que des valeurs négatives et le texte indique bien une réduction de la pression artérielle.

D : oui par définition : on a 95 % de chances de trouver la vraie valeur dans l'intervalle de confiance + 2,5 % que la différence soit supérieure à 5,50 mmHg.

E : il s'agit de la borne supérieure d'un intervalle de confiance et non d'un intervalle interquartile ; il y a donc 2,5 % de chances que la réduction soit inférieure à 5,04.

### Question 11

**Réponses exactes : A – B – C – D – E**

*Commentaires* *Cf.* figure 3 et ses commentaires.

Attention les parts de variance expliquée (mesurée par le $r^2$) sont extrêmement faibles.

L'affirmation E est vraiment sujette à caution, d'autant que rien n'est dit sur l'utilisation d'un modèle linéaire.

### Question 12

**Réponses exactes : A – B – D**

*Commentaires* A : *cf.* texte et tableaux 5.

Le p à 0,036 indique qu'il existe au moins une différence entre les quatre taux comparés deux à deux, mais ne dit pas entre quel groupe, c'est au moins vrai pour la différence de taux maximale... mais on ne peut rien dire pour les différences moins importantes (item E faux).

### Question 13

**Réponses exactes : A – B – D – E**

*Commentaires*

A et B *cf.* texte.

C : c'est la première étude.

### Question 14

**Réponses exactes : A – B – C – D**

*Commentaires*

A, B, C, D vraies = *cf.* texte.

La E est fausse c'est une hypothèse nécessaire mais c'est variabilité intra-sujets qui doit être plus petit que la variabilité inter-sujets.

### Question 15

**Réponses exactes : A – B – C – D**

*Commentaires*

Les 4 premières réponses sont justes et sont mentionnées dans la discussion (notamment page 2067, 1re colonne, 2e paragraphe.

Item E : le design est au contraire très adapté.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 10 Essai en plan factoriel

Pierre Nazeyrollas

| | |
|---|---|
| **Titre de l'article** | **Effect of Ramipril and of Rosiglitazone on Carotid Intima-Media Thickness in People With Impaired Glucose Tolerance or Impaired Fasting Glucose STARR (STudy of Atherosclerosis with Ramipril and Rosiglitazone)** |
| **Auteurs** | Lonn EM, Gerstein HC., Sheridan P, Smith S, Diaz R, Mohan V, Bosch J, Yusuf S, Dagenais GR, for the DREAM (Diabetes REduction Assessment with ramipril and rosiglitazone Medication) and STARR Investigators |
| **Référence** | J Am Coll Cardiol. 2009 ; 53(22):2028-35 |
| **Lien vers l'article** | http://www.onlinejacc.org/content/accj/53/22/2028.full.pdf |

**1** **L'étude a été réalisée car**

a. La population concernée présente un risque cardio-vasculaire élevé
b. Les glitazones permettent de diminuer l'épaisseur intima-média chez le patient présentant une intolérance au glucose
c. Les inhibiteurs de l'enzyme de conversion permettent de diminuer l'épaisseur intima-média chez le patient diabétique
d. L'association des glitazones et des inhibiteurs de l'enzyme de conversion est synergique pour diminuer le risque cardio-vasculaire
e. Les glitazones permettent de diminuer le risque cardio-vasculaire du patient diabétique

**2** **L'étude réalisée est**

a. Une étude de supériorité
b. Un essai thérapeutique
c. Un essai croisé
d. Une étude contrôlée
e. Une étude contre placebo

**3** **Pour se conformer au principe d'ambivalence, il faut que les critères de non-inclusion de l'étude comportent le ou les élément(s) suivant(s) :**

a. Présence d'une intolérance aux glitazones
b. Impossibilité de mesurer l'épaisseur intima-media
c. Présence d'une insuffisance cardiaque à fraction d'éjection altérée
d. Refus du consentement pour participer à l'étude
e. Antécédent d'angiœdème sous inhibiteur de l'enzyme de conversion



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **Le critère de jugement principal utilise dans cette étude est**

a. Défini *a posteriori*
b. Un critère de substitution
c. Composite
d. Associé au risque cardio-vasculaire
e. Reproductible

**5** **Le critère de jugement principal utilise**

a. Nécessité probablement d'inclure moins de patients qu'une analyse de la mortalité pour mettre en évidence un effet du traitement
b. Permet de comparer des moyennes
c. Permet de définir le nombre de sujets a traiter pour éviter un évènement
d. Est analysé à partir de courbes de Kaplan Meier
e. Est un critère dichotomique

**6** **Pour un risque alpha et une puissance fixés, le calcul du nombre de sujets a inclure dans l'étude dépend**

a. De la variabilité du critère de jugement
b. De la différence que l'on veut mettre en evidence
c. Du nombre de patient que l'on s'attend à perdre de vue
d. Du caractere uni ou bilatéral de l'analyse statistique
e. De la valeur attendue du critère de jugement dans le groupe de référence

**7** **Concernant la méthode de randomisation utilisée dans l'étude, on peut affirmer que**

a. Elle permet davantage de limiter les déséquilibres entre les groupes étudiés que la randomisation individuelle
b. Elle permet davantage de limiter le risque de levée de l'aveugle que la randomisation permutée par bloc
c. Elle permet de mieux équilibrer l'attribution aléatoire des traitements étudiés dans chaque centre
d. Elle permet de mieux répartir l'évolution des critères de jugement dans chaque groupe de traitement
e. Elle permet d'assurer la comparabilité des groupes de traitement tout au long de l'étude

**8** **En fonction du groupe de randomisation auquel il est affecté, un patient peut recevoir**

a. Du ramipril seul
b. Du ramipril et un placebo de rosiglitazone
c. Du ramipril et de la rosiglitazone
d. Deux placebos
e. Du ramipril et du placebo de rosiglitazone puis de la rosiglitazone et un placebo de ramipril



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**9** **L'utilisation d'un plan factoriel dans cette étude**

a. Permet d'avoir plus de puissance pour l'analyse de chaque molécule qu'en incluant le même nombre total de patient dans deux études séparées
b. Amène a comparer chaque groupe traitement au seul groupe de patients qui reçoit deux placebos
c. Permet de diminuer le biais d'attrition
d. Permet de se servir de chaque patient comme son propre témoin
e. Permet, à puissance équivalente, d'inclure moins de patients que deux études séparées

**10** **En ce qui concerne les RÉSULTATS de l'étude sur le critère de jugement principal**

a. L'interaction du ramipril et de la rosiglitazone a été analysée pour ce critère
b. La rosiglitazone a été plus efficace que le placebo
c. Le ramipril a été plus efficace que le placebo
d. La rosiglitazone a été plus efficace que le ramipril
e. Son évolution en l'absence de traitement a été deux fois moins rapide que prévu

**11** **Les analyses portant sur les objectifs secondaires montrent**

a. Une efficacité de la rosiglitazone supérieure au ramipril sur l'évolution au niveau du mur postérieur de la carotide
b. Une efficacité de la rosiglitazone supérieure au placebo sur l'évolution au niveau du mur postérieur de la carotide
c. Une efficacité de la rosiglitazone supérieure au placebo sur l'évolution au niveau du critère CIMT2
d. Une efficacité supérieure de l'association ramipril/rosiglitazone sur l'évolution au niveau du mur postérieur de la carotide par rapport au groupe double-placebo
e. L'absence de différence significative entre le ramipril et la rosiglitazone

**12** **Concernant l'adhésion au traitement en cours d'étude, il vous semble que**

a. Les patients prenant moins de 80 % du traitement prescrit sont exclus de l'analyse
b. Le mauvais suivi des traitements justifierait une analyse per protocole plutôt qu'une analyse en intention de traiter
c. Entre 10 % et 20 % des patients n'adhéraient pas correctement au traitement par placebo
d. Dans la mesure où une analyse en intention de traiter est réalisée, l'évaluation de l'adhérence au traitement ne présente pas d'interêt
e. La proportion de patients ne prenant pas correctement le traitement traduit la mauvaise conduite de cette étude

**13** **Pour les auteurs, l'absence d'efficacité du ramipril peut s'expliquer par**

a. Une sélection de patients a faible risque de progression de l'épaisseur intima-media
b. Une posologie insuffisante
c. Le bon contrôle tensionnel de la plupart des patients
d. La mauvaise tolérance du traitement propose
e. Une durée de suivi insuffisante



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**14** **Les auteurs proposent quelques arguments pouvant justifier de poursuivre les investigations. Parmi ceux-ci quels sont ceux qui pourraient conforter l'impression d'efficacité du traitement par rosiglitazone, malgré l'absence de signification sur le critère principal ?**

a. Des hypothèses physiopathologiques validées expérimentalement chez l'animal
b. Un effet significatif sur des critères secondaires validés comme étant des marqueurs de risque cardiovasculaire
c. La diminution de la mortalité cardio-vasculaire chez le diabétique
d. La diminution du risque d'insuffisance cardiaque
e. Les résultats des méta-analyses sur des critères cliniques

**15** **Au vu de cette étude etde la discussion des auteurs, quelle(s) proposition(s) semble(nt) vraie(s) concernant la prise en charge des patients intolérants au glucose ?**

a. Le ramipril et la rosiglitazone ne sont pas efficaces pour diminuer le risque vasculaire
b. Le ramipril est moins efficace que la rosiglitazone pour limiter la vitesse d'évolution de l'atteinte vasculaire
c. La prescription de rosiglitazone chez ces patients peut se justifier par précaution si le risque vasculaire est élevé
d. Les RÉSULTATS de l'étude n'amènent pas a modifier la prise en charge des patients
e. L'association des deux traitements est justifiée pour améliorer le contrôle tensionnel



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 1

**Réponse exacte : A**

*Commentaires* Tout se trouve dans le texte de l'introduction : l'étude a un intérêt pour rechercher une amélioration potentielle du risque cardio-vasculaire chez des patients à risque plus élevé que la normale.

B : si cela avait été démontré il ne serait pas nécessaire de le réévaluer à l'identique.

C : les auteurs signalent que toutes les études ne sont pas cohérentes.

D : n'est pas envisagé par les auteurs ; l'étude est construite en partant du principe qu'il n'y aura PAS d'interaction entre les deux molécules (donc pas de synergie).

E : les auteurs ne le signalent pas en introduction pour justifier l'étude, et signalent même un risque accru d'insuffisance cardiaque dans la discussion.

## Question 2

**Réponses exactes : A – B – D – E**

*Commentaires* Il n'y a pas de « piège » dans cette question... Un essai croisé correspond à une étude dans laquelle le sujet est son propre témoin et reçoit successivement (le plus souvent...) chacun des traitements. Ce n'est pas le cas ici : chaque sujet se voit attribuer par tirage au sort un seul type de traitement pour toute la durée de l'étude, parmi les 4 possibilités (association des deux principes actif, association de l'un des principes actifs et du placebo de l'autre, placebo de chaque principe actif...). C'est un plan factoriel 2x2.

## Question 3

**Réponses exactes : A – C – E**

*Commentaires* Le principe d'ambivalence implique que tout patient inclus peut recevoir n'importe lequel des traitements étudiés, donc dans ce cas le ramipril (IEC), la pioglitazone, le placebo de l'IEC ou le placebo de la glitazone. Il n'est donc pas possible d'inclure un patient qui présenterait une contre-indication à l'une des molécules (item A et E), mais il n'est pas non plus possible d'inclure un patient pour lequel l'une des molécules est nécessaire (ou une molécule de la même famille, ou un médicament avec lequel l'association serait contre-indiquée...) ; c'est le cas pour l'item C, un insuffisant cardiaque devrait être traité par un bloqueur du système rénine-angiotensine.

Les items B et C correspondent bien à des critères de non-inclusion, mais ne sont pas justifiés par le principe d'ambivalence.

Remarque : dans le contexte de l'ECN, l'item C pourrait être discuté par le jury, dans la mesure où il fait appel à des connaissances qui sont au programme mais qui sont indépendantes de l'article...

## Question 4

**Réponses exactes : B – D – E**

*Commentaires* Un critère de substitution est un critère de jugement que l'on pense associé à une critère d'évaluation clinique, mais dont l'évaluation est davantage compatible avec les délais et les effectifs disponibles. C'est le cas de l'épaisseur intima-media dans cette étude : les auteurs signalent que cette mesure est associée au risque cardio-vasculaire : « a validated marker of cardiovascular disease ». Ce critère a été défini *a priori* (item A faux), ce n'est pas un critère composite (item C faux) et une analyse de la reproductibilité de la mesure a été réalisée avec de bons résultats intra- et inter-observateurs (coefficients de corrélation intra-classes entre 0,90 et 0,96).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

■ **Question 5**

**Réponses exactes : A – B**

*Commentaires* La mortalité est faible sur une durée de suivi relativement courte (3 décès parmi plus de 1 400 patients...) : il semble effectivement impossible de mettre en évidence un effet du traitement sur la mortalité dans cette étude. Le critère de jugement est une variable continue, ce n'est donc pas un critère dichotomique, il n'y a pas de définition « d'évènement « ni d'analyse de « survie ».

■ **Question 6**

**Réponses exactes : A – B – C – D – E**

*Commentaires*

- Plus la variabilité de la mesure du critère de jugement est grande, plus il faudra de sujets.
- Plus la différence à mettre en évidence est petite, plus il faudra de sujets.
- Dans une étude longitudinale, il est possible de perdre de vue des sujets lors du suivi. Dans ce cas, il faudrait anticiper cette perte et la rajouter au nombre de sujets à inclure.
- Avec une hypothèse unilatérale, on a besoin de tester qu'un seul sens. Donc on aura besoin de moins de sujets qu'avec une hypothèse bilatérale.
- La valeur attendue dans le groupe de référence influence la différence à mettre en évidence.

■ **Question 7**

**Réponses exactes : A – C**

*Commentaires*

- La randomisation par bloc consiste à réaliser un tirage au sort sur des blocs de malades (d'effectif toujours pair), permettant d'avoir à la fin de chaque bloc une égale répartition des sujets entre les groupes de traitement. Par exemple, une randomisation par bloc de 8 assure que tous les 8 sujets inclus, 2 auront été randomisés dans chacun des 4 groupes de traitement. Donc, elle permet, par rapport à la randomisation individuelle de limiter les déséquilibres (réponse A vraie).
- L'inconvénient de la randomisation par bloc est l'anticipation possible, pour le dernier malade de chaque bloc, du traitement qui va lui être attribué. C'est pour cela que dans certains cas, la taille des blocs varie au cours de l'étude : c'est la randomisation permutée par bloc (réponse B fausse).
- Ici la randomisation a été stratifiée par centre afin d'assurer l'équilibre de répartition des facteurs de confusion potentiels au sein des groupes de traitement dans chaque centre (réponse C vraie).
- La randomisation ne peut permettre d'assurer la répartition de l'évolution des critères de jugement dans chaque groupe, car ceux-ci lui sont postérieurs (réponse D fausse).
- De même, après la randomisation les groupes peuvent changer (sortie d'étude par exemple), déséquilibrant ainsi les groupes (réponse E fausse).

■ **Question 8**

**Réponses exactes : B – C – D**

*Commentaires* Plan factoriel 2x2 avec le ramipril, la rosiglitazone et un placebo pour chaque molécule : 4 groupes sont définis, un patient ne change pas de groupe en cours d'étude.

| | Ramipril | Placebo de Ra |
|---|---|---|
| Rosiglitazone | Ra+Ro | pcbRa+Ro |
| Placebo de Ro | Ra+pcbRo | pcbRa+pcbRo |



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

■ **Question 9**

**Réponses exactes : A – E**

*Commentaires*

Les items A et E sont équivalents : le plan factoriel permet dans ce cas d'analyser l'effet de chaque molécule (ramipril ou rosiglitazone) par rapport à l'ensemble des patients recevant le placebo correspondant en prenant en compte l'effet éventuel de l'autre molécule. Cette analyse est pertinente en l'absence d'interaction entre les deux molécules testées, ce que les auteurs ont vérifié.

B : le groupe de patients recevant un traitement n'est pas comparé au seul groupe recevant les deux placebos, sauf si une interaction entre les traitements est mise en évidence, ce qui n'est pas le cas ici.

C : ce type d'étude n'a aucun rapport avec le biais d'attrition (c'est un biais du à l'exclusion de patients au cours de l'étude).

D : Le patient est son propre témoin dans le cadre d'un essai croisé, ce qui n'est pas le cas ici,

■ **Question 10**

**Réponses exactes : A – E**

*Commentaires* Les réponses sont dans le texte...

A : « ... after confirming that there was no significant interaction between the study treatments for the primary or secondary CUS outcomes... » l'analyse a été faite et il n'y a pas d'interaction significative (p = 0,36).

B : tableau 3, p = 0,08

C : tableau 3, p = 0,37

D : les molécules n'ont pas été comparées l'une à l'autre, ce n'est pas l'objet de l'étude.

E : les auteurs avaient prévu une progression de 0,016 mm/an sous placebo pour évaluer le nombre de sujets à inclure, la progression dans le groupe recevant deux placebos a été de 0,0076 mm/an (dans le texte), ce qui est bien environ deux fois moins que prévu.

■ **Question 11**

**Réponses exactes : B – C**

*Commentaires* Pour les item B et C la réponse se trouve dans le tableau 3 (avec pour chacun p = 0,01).

Il n'y a pas eu de comparaison de la rosiglitazone avec le ramipril (item A et item E) : on ne peut donc pas conclure sur la supériorité d'une molécule par rapport à l'autre, ni sur l'absence de différence...

De la même façon, les patients recevant l'association des deux molécules n'ont pas été comparés à ceux recevant deux placebos (item D).

■ **Question 12**

**Réponse exacte : C**

*Commentaires* Les auteurs définissent l'adhésion au traitement par la prise d'au moins 80 % du traitement prescrit ; en revanche les patients qui prennent mal leur traitement ne sont pas exclus de l'analyse qui est réalisée en intention de traiter. Dans une étude de supériorité, comme c'est le cas ici, l'analyse se fait en intention de traiter ce qui permet de minimiser les biais liés à des exclusions survenues après la randomisation.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Avoir une adhésion au traitement de l'ordre de 80 % n'est pas anormalement bas dans un essai thérapeutique de longue durée : il ne s'agit pas d'un mauvais suivi du traitement et une mauvaise observance n'est pas une bonne raison pour ne pas réaliser l'analyse en intention de traiter.

L'évaluation de l'adhésion au traitement présente un intérêt, quel que soit le type d'analyse réalisé : c'est un élément qui permet d'apprécier la tolérance, de dépister des effets indésirables inattendus ou qui peut se révéler utile lors de la discussion.

La proportion de patients « mal-adhérents » constatée n'est pas inhabituelle et ne reflète pas ici une mauvaise conduite de l'étude. Le fait que l'adhésion ait pu être analysée est plutôt un marqueur de bonne qualité méthodologique.

L'item C est juste : la proportion de patients adhérent au traitement est donnée par les auteurs dans le texte.

## Question 13

**Réponses exactes : A – C – E**

*Commentaires* Les auteurs se disent « surpris » par l'absence d'effet significatif du ramipril. Les arguments pour l'expliquer sont :

- des patients qui ne sont pas à haut risque de progression comme le montre la faible évolution de l'EIM par rapport à ce qui était attendu (item A) ;
- une durée d'étude courte (item E). Cet élément est présenté par les auteurs mais peut être discutable puisque l'analyse principale porte sur une mesure de progression annuelle dont rien n'indique qu'elle se modifie lors d'un suivi plus long ;
- les auteurs signalent que les patients avaient déjà pour la plupart un bon contrôle tensionnel, limitant l'effet sur la plaque que pourrait avoir le ramipril en faisant baisser la tension artérielle (item C) ;
- les auteurs ne signalent pas une posologie insuffisante (les posologies de ramipril utilisées sont habituelles, voire plutôt élevées après un an). Même si la tolérance était mauvaise (ce qui n'est pas le cas, et les auteurs ne le signalent pas), l'efficacité d'un traitement dépend de la possibilité de l'administrer : il serait donc difficile de justifier un manque d'efficacité par une mauvaise tolérance.

## Question 14

**Réponses exactes : A – B**

*Commentaires*

Les hypothèses physiopathologiques et les données expérimentales, même chez l'animal, peuvent être des éléments de validité externe : les auteurs signalent un effet expérimental vasculo-protecteur des glitazones et une diminution de l'athérome lors de l'administration de rosiglitazone chez l'animal (item A).

Ils signalent également d'autres études montrant une diminution de la progression de l'épaisseur intima-media (item B).

Il n'y a en revanche pas de données de morbi-mortalité chez le diabétique (item C) ; le risque d'insuffisance cardiaque est augmenté (et non diminué, item D) ; et les méta-analyses montrent une augmentation du risque cardio-vasculaire ce qui n'est évidemment pas en faveur de l'efficacité des glitazones (item E).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 15**

**Réponse exacte : D**

*Commentaires*

Le risque vasculaire des patients n'est pas analysé : ce qui l'est ce sont des mesures d'épaisseur de la paroi carotidienne : aucune conclusion concernant le risque n'est possible.

Le ramipril et la rosiglitazone ne sont pas comparés... (bis !).

L'étude ne présente aucun élément pouvant faire retenir un traitement par rosiglitazone devant un risque vasculaire élevé, les données physiopathologiques et animales évoquées dans la discussion ne sont pas suffisantes et les données cliniques évoquées sont en faveur d'une augmentation du risque !

Le faible effet hypotenseur signalé par les auteurs pour la rosiglitazone ne justifie pas son utilisation dans cette indication...

La seule réponse adéquate est D.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 11 Essai d'intervention

José Labarère

| | |
|---|---|
| **Titre de l'article** | **Multifaceted intervention to prevent venous thromboembolism in patients hospitalized for acute medical illness : a multicenter cluster-randomized trial** |
| **Auteurs** | Roy P-M, Rachas A, Meyer G, Le Gal G, Durieux P, El Kouri D, Honnart D, Schmidt J, Legall C, Hausfater P, Chrétien J-M, Mottier D, and PREVENU study group |
| **Référence** | PLoS ONE 11(5): e0154832. |
| **Lien vers l'article** | **http://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0154832&type=printable** |

**1** **Parmi les motifs d'admission à l'hôpital suivants, quels sont ceux qui constituaient des critères d'exclusion de l'essai ?**

a. Embolie pulmonaire massive avec choc à la prise en charge par le SMUR
b. Infarctus du myocarde avec sus-décalage du segment ST
c. Décompensation d'une insuffisance cardiaque secondaire à un épisode d'arythmie complète par fibrillation auriculaire
d. Pneumonie aiguë communautaire sans facteur de risque de maladie thrombo-embolique veineuse associé
e. Thrombose veineuse proximale du membre inférieur gauche au troisième trimestre de grossesse

**2** **L'unité de randomisation de l'essai est**

a. Le patient
b. Le médecin ou l'interne du service d'accueil des urgences
c. La période d'étude (pré-intervention *versus* intervention)
d. Le centre hospitalier
e. L'intervention multifacette

**3** **Dans cet essai, l'unité de randomisation est le centre hospitalier. Par comparaison à la randomisation individuelle des patients, la randomisation des centres hospitaliers**

a. Limite le risque de contamination du bras contrôle par le bras intervention
b. Se justifie par le fait que l'intervention multifacette cible les pratiques médicales de prescription
c. Constitue des bras comparables pour les caractéristiques à l'inclusion des patients
d. Viole l'hypothèse d'indépendance des observations
e. Réduit les fluctuations aléatoires d'échantillonnage



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **Dans cet essai, le calcul du nombre de sujets nécessaire reposait sur**

a. Les capacités de recrutement de chaque centre hospitalier
b. Un coefficient de corrélation intra-cluster de 0,01
c. Une puissance statistique de 80 %
d. L'hypothèse d'une incidence cumulée attendue du critère de jugement principal à 90 jours de 5 % dans le bras contrôle
e. Une réduction absolue attendue de 1,5 point de pourcentage du critère de jugement principal à 90 jours dans le bras intervention

**5** **La proportion de séjours de durée supérieure à 14 jours était significativement plus élevée dans le bras contrôle que dans le bras intervention de l'essai (1775/9071 [19,6 %] *versus* 1 791/7682 [23,3 %], p < 0,001, p-valeur calculée d'après le tableau 2). Qu'est-ce qui a pu contribuer à cette différence observée ?**

a. Une fréquence plus élevée d'événements thrombo-emboliques veineux symptomatiques dans le bras contrôle
b. Le risque d'erreur statistique de première espèce (alpha)
c. Une prévalence plus élevée de l'insuffisance cardiaque congestive dans le bras contrôle
d. Une durée moyenne de séjour globalement plus longue dans les hôpitaux du bras contrôle, à recrutement comparable
e. Un défaut de puissance statistique

**6** **Les éléments de l'intervention multifacette qui ont pu être implémentés dans les 13 centres du bras intervention de l'essai incluaient**

a. La tenue de conférences pédagogiques
b. La mise à disposition d'une application pour téléphone mobile
c. Un logiciel de rappels au moment de la décision
d. La réalisation d'un audit avec retour d'information
e. La diffusion de matériel pédagogique (affiches, fiches au format de poche)

**7** **Quels sont les deux événements qui expliquent la part la plus importante de l'incidence cumulée du critère de jugement principal composite ?**

a. L'embolie pulmonaire, fatale ou non
b. La mort subite inexpliquée
c. La thrombose veineuse profonde distale
d. L'événement hémorragique majeur, fatal ou non
e. La mortalité toute cause



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**8** **Quels éléments manquent pour s'assurer de la validité de l'évaluation du critère de jugement principal ?**

a. La recherche de thromboses veineuses profondes asymptomatiques par un examen écho-doppler veineux systématique
b. L'adjudication des événements par un comité indépendant en aveugle du bras de l'essai
c. La confirmation des événements thrombo-emboliques veineux symptomatiques par un test diagnostique objectif
d. La confirmation des embolies pulmonaires et des hémorragies fatales par une autopsie
e. La description des modalités de recueil des suspicions d'événements cliniques pendant les 3 mois de suivi

**9** **Quelles informations contenues dans le tableau 3 permettent de conclure que l'étude n'a pas mis en évidence de différence significative de l'incidence cumulée du critère de jugement principal entre les deux bras de l'essai ?**

a. Le degré de signification associé au bras de l'essai était de 0,97, après ajustement sur le cluster
b. Le degré de signification associé au bras de l'essai était de 0,87, après ajustement sur le cluster et les facteurs de confusion potentiels
c. L'estimation ponctuelle de l'odds ratio du critère de jugement principal ajusté sur le cluster (0,99) était comprise entre les bornes de son intervalle de confiance à 95 % (0,75 à 1,32)
d. L'intervalle de confiance à 95 % de l'odds ratio du critère de jugement principal ajusté sur le cluster (0,75 à 1,32) incluait la valeur neutre 1
e. L'intervalle de confiance à 95 % de l'odds ratio du critère de jugement principal ajusté sur le cluster (0,75 à 1,32) incluait la valeur neutre 0

**10** **Quelles informations contenues dans le tableau 4 permettent de conclure que l'étude n'a pas mis en évidence de différence significative de l'évolution du pourcentage de prophylaxie pharmacologique adéquate de la maladie thrombo-embolique veineuse entre les deux bras de l'essai ?**

a. Le degré de signification associé au terme d'interaction entre le bras et la période de l'essai était de 0,65
b. La différence entre les deux bras de l'essai de la variation du pourcentage de prophylaxie pharmacologique adéquate était ajustée sur les facteurs de confusion potentiels
c. L'estimation ponctuelle de la différence ajustée de la variation de prophylaxie pharmacologique adéquate (-1,2 %) était comprise entre les bornes de son intervalle de confiance à 95 % (-6,6 à 4,2)
d. L'intervalle de confiance à 95 % de la différence ajustée de la variation de prophylaxie pharmacologique adéquate (-6,6 à 4,2) incluait la valeur neutre 1
e. L'intervalle de confiance à 95 % de la différence ajustée de la variation de prophylaxie pharmacologique adéquate (-6,6 à 4,2) incluait la valeur neutre 0



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**11** **Quelles sont les explications potentielles à l'absence de différence significative de l'évolution du pourcentage de prophylaxie pharmacologique adéquate de la maladie thrombo-embolique veineuse entre les deux bras de l'essai ?**

a. Le déséquilibre de l'effectif de patients entre les deux périodes, au sein de chaque bras de l'essai
b. Le risque d'erreur statistique de première espèce (alpha)
c. L'absence d'impact de l'intervention multifacette sur les pratiques de prophylaxie pharmacologique de la maladie thrombo-embolique veineuse
d. Une modification des pratiques de prophylaxie pharmacologique de la maladie thrombo-embolique veineuse dans le bras contrôle
e. La fenêtre temporelle d'observation retenue pour qualifier l'adéquation de la prophylaxie pharmacologique

**12** **Pendant la période d'intervention, le pourcentage de patients recevant une prophylaxie pharmacologique était de 48,2 % dans le bras « intervention multifacette » *versus* 44,1 % dans le bras contrôle (P < 0,001). En dépit de cet écart, l'essai n'a pas mis en évidence de différence de l'incidence cumulée d'événements thrombo-emboliques veineux entre les deux bras de l'essai (1,9 % *versus* 1,9 %) sur la même période. Cette observation paradoxale pourrait s'expliquer par**

a. Une inefficacité de la prophylaxie pharmacologique des événements thrombo-emboliques veineux
b. Une hétérogénéité des pratiques de prophylaxie mécanique entre les deux bras de l'essai
c. Une incidence cumulée d'événements thrombo-emboliques veineux plus faible qu'attendue dans le bras contrôle
d. Un écart d'amplitude insuffisant de la fréquence d'utilisation de la prophylaxie pharmacologique pour détecter une différence significative de l'incidence d'événements thrombo-emboliques veineux entre les deux bras de l'essai
e. Un pourcentage plus élevé de patients à haut risque d'événement thrombo-embolique veineux dans le bras « intervention multifacette »

**13** **Cette étude**

a. Démontre la supériorité de l'intervention multifacette, comparativement à l'absence d'intervention, pour réduire l'incidence cumulée du critère de jugement principal
b. Démontre l'équivalence de l'intervention multifacette, comparativement à l'absence d'intervention, pour réduire l'incidence cumulée du critère de jugement principal
c. Démontre la non-infériorité de l'intervention multifacette, comparativement à l'absence d'intervention, pour réduire l'incidence cumulée du critère de jugement principal
d. Ne met pas en évidence d'association significative entre le bras de l'essai et l'incidence cumulée du critère de jugement principal
e. Démontre l'absence d'association significative entre l'adéquation de la prophylaxie pharmacologique de la maladie thrombo-embolique veineuse et l'incidence cumulée du critère de jugement principal



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**14** **Quelles raisons pourraient expliquer le fait que l'adéquation de la prophylaxie pharmacologique de la maladie thrombo-embolique veineuse n'ait pas progressé dans le bras « intervention multifacette » ?**

a. La validité discutable des recommandations utilisées
b. La complexité de l'évaluation du risque thrombo-embolique veineux en routine
c. L'absence de définition standardisée des facteurs de risque de maladie thrombo-embolique veineuse
d. L'implantation incomplète de l'intervention multifacette dans 11 des 13 centres hospitaliers participants
e. L'inefficacité des rappels informatisés au moment de la décision pour améliorer les pratiques de prophylaxie pharmacologique de la maladie thrombo-embolique veineuse

**15** **Quelles sont les limites de cet essai ?**

a. La sur-représentation des centres hospitaliers universitaires parmi les centres participants
b. L'utilisation de recommandations pour la prophylaxie de la maladie thrombo-embolique veineuse datant de 2008
c. L'impossibilité d'implémenter les logiciels d'aide à la prescription et les rappels informatisés au moment de la décision dans la majorité des centres hospitaliers
d. Le recrutement des centres hospitaliers en France uniquement
e. L'absence de description des molécules utilisées en prophylaxie



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 1

**Réponses exactes : A – B – C – E**

*Commentaires* Les épisodes de maladie thrombo-embolique veineuse présents à l'admission (ou dans les 48 heures suivant l'admission) sont des critères d'exclusion de l'étude (réponses A et E vraies, page 3). Les pathologies justifiant un traitement anticoagulant autres que la maladie thrombo-embolique veineuse sont également un critère d'exclusion (réponse B et C vraies, page 3). En revanche, l'admission pour une pneumonie aiguë communautaire sans facteur de risque de maladie thrombo-embolique veineuse associé ne constitue pas un critère d'exclusion de l'étude (réponse D fausse) : il ne faut pas confondre critère d'exclusion de l'essai et non-indication de la prophylaxie pharmacologique.

### Question 2

**Réponse exacte : D**

*Commentaires* L'unité de randomisation de cet essai d'intervention en cluster est le centre hospitalier (paragraphe intervention period, page 3, réponse D vraie).

### Question 3

**Réponses exactes : A – B – D**

*Commentaires* Le niveau pertinent d'implantation de l'intervention multifacette était l'hôpital car celle-ci ciblait les pratiques médicales de prévention de la maladie thrombo-embolique veineuse et les infrastructures informatiques (logiciel d'aide à la décision clinique et rappel au moment de la décision) (réponse B vraie). Dans un essai avec une randomisation individuelle, il est impossible de demander au médecin de s'affranchir de l'effet de l'intervention lorsqu'il prend en charge un patient du bras contrôle. En randomisant les centres hospitaliers, on évite cette contamination du bras contrôle par le bras intervention : tous les patients d'un centre hospitalier sont pris en charge par des médecins exposés soit à l'intervention multifacette soit à l'intervention contrôle (réponse A vraie). La randomisation par cluster risque davantage de constituer des groupes déséquilibrés à l'aune des caractéristiques à l'inclusion des patients, reflétant l'hétérogénéité de recrutement entre les hôpitaux (réponse C fausse). Les patients d'un centre hospitalier sont davantage semblables entre eux que semblables aux patients des autres centres hospitaliers : la randomisation par cluster viole l'hypothèse d'indépendance des observations (réponse D vraie). La randomisation (par cluster ou individuelle) est soumise aux fluctuations aléatoires d'échantillonnage (réponse E fausse).

### Question 4

**Réponses exactes : B – C – D – E**

*Commentaires* Le calcul du nombre de sujets nécessaire dans un essai de supériorité pour un critère de jugement qualitatif binaire est basé sur le risque d'erreur statistique de première espèce alpha consenti, la puissance statistique souhaitée (réponse C vraie), l'hypothèse sur la fréquence du critère de jugement principal dans le groupe contrôle (réponse D vraie), et la taille de l'effet (réponse E vraie). Le calcul du nombre de sujets nécessaire n'implique pas les capacités de recrutement des centres participants (réponse A fausse) : les capacités de recrutement donnent une indication sur les possibilités d'atteindre le nombre de sujets nécessaire calculé. Dans un essai randomisé par clusters, l'hypothèse d'indépendance des observations est violée : le calcul du nombre de sujets nécessaire doit intégrer cette corrélation des obser-



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

vations au sein de chaque cluster (réponse B vraie). Le nombre de sujets nécessaire d'un essai randomisé en cluster est égal au nombre de sujets nécessaire d'un essai randomisé individuel multiplié par un facteur d'inflation.

### Question 5

**Réponses exactes : B – C – D**

*Commentaires* L'étude n'a justement pas mis en évidence de différence significative de l'incidence cumulée d'événements thrombo-emboliques veineux entre les deux bras de l'essai (réponse A fausse). La proportion plus élevée de séjours de durée supérieure à 14 jours dans le bras contrôle peut résulter de fluctuations d'échantillonnage (réponse B vraie), de la non-comparabilité des caractéristiques à l'admission des patients (réponse C vraie), ou refléter des durées moyennes de séjour plus longues à recrutement comparable (réponse D vraie). Une différence statistiquement significative ne peut jamais être le fruit d'un défaut de puissance statistique (réponse E fausse).

### Question 6

**Réponses exactes : A – E**

*Commentaires* Tous les centres du bras intervention de l'essai ont bénéficié de la tenue de conférences pédagogiques (réponse A), de la diffusion de matériel pédagogique (réponse E), et de la mise à disposition d'un logiciel téléchargeable d'aide à la décision clinique (pages 3 et 6). En revanche, seuls deux centres du bras intervention ont pu effectivement installer le logiciel de rappel au moment de la décision (réponse C fausse, page 6). Le protocole de l'étude ne prévoyait pas d'audit avec retour d'information (réponse D fausse).

### Question 7

**Réponses exactes : B – D**

*Commentaires* La mort subite inexpliquée et l'événement hémorragique majeur, fatal ou non, sont à l'origine de 129 (28 %) et 186 (41 %) des 464 cas du critère de jugement principal (Tableau 3, réponses B et D vraies). L'embolie pulmonaire, fatale ou non, et la thrombose veineuse profonde distale sont à l'origine de 67 (14 %) et 36 (8 %) des 464 cas du critère de jugement principal (Tableau 3, réponses A et C fausses). La mortalité toute cause ne participe pas au critère de jugement principal (réponse E fausse) : seuls les décès inexpliqués, les embolies pulmonaires fatales et les événements hémorragiques fatals participent au critère de jugement principal composite.

### Question 8

**Réponses exactes : D – E**

*Commentaires* Les thromboses veineuses profondes asymptomatiques ne sont pas forcément cliniquement pertinentes, notamment lorsqu'elles sont distales (réponse A fausse). L'adjudication des événements a été réalisée par un comité indépendant, en aveugle du bras de l'essai (réponse B fausse) et les événements thrombo-emboliques veineux ont été confirmés par un test diagnostique. En revanche, l'article n'indique pas si les embolies pulmonaires ou les hémorragies fatales ont été documentées par autopsie (réponse D vraie). Il ne décrit pas non plus comment ont été recueillies les suspicions d'événement cliniques pendant les 3 mois de suivi (réponse E vraie).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 9

**Réponses exactes : A – B – D**

*Commentaires* Les éléments permettant de conclure que l'étude n'a pas mis en évidence de différence significative de l'incidence cumulée du critère de jugement principal entre les deux bras de l'essai étaient le degré de signification du test statistique (réponses A et B vraies) et l'intervalle de confiance de l'odds ratio. La valeur neutre de l'odds ratio est de 1 (modèle multiplicatif, réponse D vraie). La valeur neutre 0 serait pertinente pour un modèle additif (différence de risque, réponse E fausse). Il est normal que l'estimation ponctuelle soit encadrée par les bornes de son intervalle de confiance à 95 % : cette information ne permet pas de conclure à une différence significative (réponse C fausse).

### Question 10

**Réponses exactes : A – E**

*Commentaires* Les éléments permettant de conclure que l'étude n'a pas mis en évidence de différence significative de l'évolution du pourcentage de prophylaxie pharmacologique adéquate de la maladie thrombo-embolique veineuse entre les deux bras de l'essai étaient le degré de signification du test d'interaction (réponse A vraie) et l'intervalle de confiance de la différence. La valeur neutre de cette différence est de 0 (réponse E vraie). La valeur neutre 1 serait pertinente pour un modèle multiplicatif (odds ratio ou risque relatif, réponse D fausse). Il est normal que l'estimation ponctuelle soit encadrée par les bornes de son intervalle de confiance à 95 % : cette information ne permet pas de conclure à une différence significative (réponse C fausse). L'ajustement vise à contrôler les facteurs de confusion potentiels mais ne permet pas de conclure sur la signification statistique de la différence (réponse B fausse).

### Question 11

**Réponses exactes : C – D**

*Commentaires* Les raisons qui pourraient expliquer l'absence de différence significative de l'évolution du pourcentage de prophylaxie pharmacologique adéquate entre les deux bras de l'essai sont l'absence d'impact de l'intervention multifacette sur les pratiques médicales (réponse C vraie) et un effet Hawthorne dans le bras contrôle (réponse D vraie). Le risque d'erreur statistique de première espèce (alpha) ne peut pas expliquer l'absence de différence significative sur un échantillon (réponse B fausse). Une analyse de sensibilité montre que changer la fenêtre temporelle d'observation (2 jours *versus* 5 jours) pour qualifier l'adéquation de la thromboprophylaxie ne modifiait pas les résultats (page 9, réponse E fausse).

### Question 12

**Réponses exactes : B – C – D**

*Commentaires* Les raisons potentielles à l'origine de l'absence de différence significative de l'incidence cumulée d'événements thrombo-emboliques veineux entre les deux bras de l'essai incluent l'existence de facteurs de confusion, notamment une hétérogénéité des traitements associés (réponse B vraie), une augmentation insuffisante de la fréquence d'utilisation de la prophylaxie pharmacologique pour prévenir un nombre supplémentaire d'événements thrombo-emboliques veineux dans le bras « intervention multifacette » (réponse D vraie), et une fréquence observée d'événements thrombo-emboliques veineux plus faible que celle attendue dans le groupe contrôle (réponse C vraie). L'inefficacité théorique de la prophylaxie pharmacologique de la maladie thrombo-embolique veineuse est réfutée par les méta-analyses



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

d'essais randomisés (référence 4 citée en introduction, réponse A fausse). La prévalence du risque thrombo-embolique veineux élevé était de 30,7 % dans le bras « intervention multifacette » *versus* 33,0 % dans le bras contrôle (Tableau 2, réponse E fausse).

### Question 13

**Réponse exacte : D**

*Commentaires* L'essai est négatif : il n'a pas mis en évidence de différence significative de l'incidence cumulée du critère de jugement principal entre le bas d'intervention multifacette et le bras contrôle. Une autre formulation acceptable est que l'étude ne met pas en évidence d'association significative entre le bras de l'essai et l'incidence cumulée du critère de jugement principal (réponse D vraie et réponse A fausse). Le fait que l'essai ne mette pas en évidence de différence significative du critère de jugement principal entre les deux bras n'autorise pas à conclure à l'équivalence ou à la non-infériorité de l'intervention multifacette comparativement à l'absence d'intervention (réponses B et C fausses). Pour conclure à la non-infériorité de l'intervention multifacette, il aurait été nécessaire de définir *a priori* une marge de non-infériorité de l'incidence cumulée du critère de jugement principal. En revanche, l'essai n'a pas été planifié pour tester l'existence d'une association entre l'adéquation de la prophylaxie pharmacologique et l'incidence cumulée du critère de jugement principal – l'article ne rapporte pas non plus une telle analyse (réponse E fausse).

### Question 14

**Réponses exactes : A – B – C – D**

*Commentaires* Les raisons potentielles du résultat négatif de l'essai pour les pratiques de prophylaxie pharmacologique de la maladie thrombo-embolique veineuse sont citées en discussion de l'article (pages 10 et 11) : l'implantation incomplète de l'intervention multifacette (seuls deux centres ont implémenté les rappels informatisés au moment de la décision, réponse D vraie), la complexité de l'évaluation du risque thrombo-embolique veineux en routine (réponse B vraie), l'absence de définition standardisée des facteurs de risque de maladie thrombo-embolique veineuse (réponse C vraie). La validité discutable des recommandations utilisées datant de 2008 pourrait également expliquer la faible observance des médecins, bien que la date de réalisation de l'essai ne soit pas indiquée (réponse A). En revanche, les rappels au moment de la décision et les logiciels d'aide à la prescription ont démontré leur efficacité pour modifier les pratiques de prophylaxie de la maladie thrombo-embolique veineuse (référence 20 citée en discussion, réponse E fausse).

### Question 15

**Réponses exactes : A – B – C – D – E**

*Commentaires* Les limites en termes de validité externe sont la sur-représentation des centres hospitaliers universitaires (réponse A vraie), le recrutement des centres hospitaliers uniquement en France (réponse D vraie), l'utilisation de recommandations qui ont depuis été actualisées en partie (réponse B vraie), et l'absence d'information sur les molécules ou classes thérapeutiques utilisées (héparine de bas poids moléculaire, héparine non fractionnée, anticoagulants oraux, nouveaux anticoagulants, réponse E vraie). Finalement, l'implantation incomplète de l'intervention multifacette dans la plupart des centres limite les possibilités d'interprétation du résultat négatif de l'essai (réponse E vraie).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 12 Essai d'intervention

Antoine Duclos, Pauline Occeli

| | |
|---|---|
| **Titre de l'article** | **Multifaceted intervention to decrease the rate of severe post-partum haemorrhage : the PITHAGORE6 cluster-randomised controlled trial** |
| **Auteurs** | Deneux-Tharaux C, Dupont C, Colin C, Rabilloud M, Touzet S, Lansac J, Harvey T, Tessier V, Chauleur C, Pennehouat G, Morin X, Bouvier-Colle MH, Rudigoz R |
| **Référence** | Br J Obstet Gynecol. 2010 Sep ; 117(10):1278-87 |
| **Lien vers l'article** | **http://onlinelibrary.wiley.com/doi/10.1111/j.1471-0528.2010.02648.x/epdf** |

**1** **D'après vous, l'hypothèse principale testée dans cette étude était**

a. Qu'un programme multifacette destiné à améliorer la prise en charge précoce de l'hémorragie du post-partum pourrait en réduire la survenue
b. Qu'une formation pourrait réduire la survenue d'hémorragies du post-partum
c. Qu'un programme multifacette pourrait améliorer le pronostic des femmes présentant une hémorragie sévère du post-partum
d. Qu'un programme multifacette destiné à améliorer les pratiques de prise en charge précoce de l'hémorragie du post-partum pourrait réduire la survenue d'hémorragies sévères du post-partum
e. Qu'une formation destinée à améliorer la prise en charge de l'accouchement pourrait en réduire la morbi-mortalité

**2** **Le schéma d'étude proposé par les auteurs intégrait**

a. Un groupe contrôle
b. Une comparaison du critère de jugement avant-après l'intervention multifacette
c. Une randomisation individuelle des femmes à l'accouchement
d. Plusieurs maternités réparties dans plusieurs réseaux de périnatalité
e. Deux bras en parallèle

**3** **Le « cluster » décrit dans le papier**

a. Correspondait à l'ensemble des patientes prises en charge durant l'étude
b. Correspondait à chaque maternité participant à l'étude
c. Correspondait à chaque réseau de périnatalité participant à l'étude
d. Était lié aux maternités appartenant exclusivement au groupe intervention de l'étude
e. Était justifié pour limiter le risque de contamination du groupe contrôle par l'intervention



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **L'intervention évaluée dans l'étude**

a. Comprenait une analyse de cas d'hémorragies sévères du post-partum survenus localement
b. A été mise en place dans l'ensemble des maternités en deux phases successives
c. A été conduite par un binôme « obstétricien/sage-femme » dans chaque maternité du groupe intervention
d. Reposait sur la combinaison de trois éléments
e. Ciblait directement les patientes prises en charge durant l'étude

**5** **Le critère de jugement principal choisi par les auteurs**

a. Correspondait à un critère de résultat clinique
b. Mesurait la survenue d'hémorragies du post-partum
c. S'exprimait sous la forme d'un taux
d. Était constitué du regroupement de plusieurs critères
e. Provenait d'une extraction automatique à partir du dossier informatisé des patientes

**6** **La randomisation des maternités**

a. Était stratifiée en fonction des caractéristiques des femmes prises en charge
b. Visait à permettre la comparabilité des maternités entre les deux bras selon leur réseau périnatal d'appartenance et leur volume d'accouchement
c. Visait à garantir la comparabilité initiale du critère de jugement principal entre les deux bras de l'essai
d. Était réalisée de manière centralisée au moment de l'inclusion de chaque nouvelle patiente dans l'essai
e. Permettait de tirer au sort les patientes collectivement en fonction de leur maternité de prise en charge

**7** **Le nombre de sujets nécessaire a été calculé**

a. En intégrant un facteur d'inflation reposant sur le coefficient de corrélation intracluster
b. Pour garantir une puissance statistique de 80 %
c. Pour objectiver une différence absolue de 0,6 % sur le critère principal entre les deux bras de l'essai
d. En faisant l'hypothèse qu'au moins 1 340 femmes seraient incluses dans chacune des maternités
e. En supposant un taux initial d'hémorragies sévères du post-partum de 1 %

**8** **La population finalement analysée**

a. Comprenait 146 781 accouchements réalisés dans 106 maternités
b. Comprenait des maternités avec des volumes annuels d'accouchement égaux
c. Présentait une taille suffisante compte tenu des hypothèses faites pour calculer le nombre de sujets nécessaire
d. Reposait sur une approche de type « per protocole »
e. Était représentative de l'ensemble des accouchements réalisés dans les maternités françaises



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**9** **L'interprétation des résultats suggérait**

a. La comparabilité initiale des deux bras de l'étude sur le critère de jugement principal
b. L'efficacité de l'intervention testée sur le critère de jugement principal
c. L'application plus fréquente de certaines pratiques recommandées dans le bras intervention comparativement au bras contrôle
d. Une proportion différente d'accouchements par césarienne entre les deux bras
e. L'influence du type de maternité sur la survenue d'hémorragies du post-partum

**10** **Concernant la prise en charge des hémorragies du post-partum**

a. L'administration de sulprostone dans les 30 minutes était statistiquement plus élevée dans le bras intervention
b. L'appel à l'aide d'un collègue senior était plus fréquent dans le bras intervention
c. L'appel à l'aide d'un collègue senior réduisait par un facteur 1,65 le risque de survenue d'une hémorragie sévère
d. La réalisation d'une révision de la cavité utérine ne différait pas entre les deux bras
e. Aucune amélioration n'était observée sur l'ensemble des critères dans le bras intervention

**11** **L'évolution trimestrielle de la proportion d'hémorragies sévères du post-partum**

a. Démontrait une amélioration statistiquement significative dans les deux bras de l'essai
b. Était en faveur d'un effet de l'intervention testée
c. Objectivait une différence de tendance entre les deux groupes
d. N'intégrait pas la tendance préexistante à la mise en place de l'intervention
e. Apportait la preuve d'un biais de contamination entre les deux bras de l'essai

**12** **L'absence de différence entre les deux bras de l'essai sur le critère principal pourrait s'expliquer par**

a. Un possible déséquilibre initial entre les bras de l'étude sur le critère principal
b. L'absence d'effet de l'intervention
c. L'absence d'insu des maternités de leur appartenance au groupe intervention ou contrôle de l'étude
d. Un éventuel biais de contamination entre les bras
e. Une tendance à l'amélioration des résultats cliniques au cours du temps

**13** **Cette étude a été conduite et publiée**

a. Par un groupe d'investigateurs appartenant notamment à différents réseaux de périnatalité français
b. Après déclaration auprès de la Commission Nationale de l'Informatique et des Libertés
c. Avec l'accord d'un Comité de Protection des Personnes
d. Sans financent particulier
e. Dans une revue scientifique de pédiatrie



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**14** **Quelles sont les caractéristiques méthodologiques principales de cette étude ?**

a. En insu
b. En cluster
c. Randomisée
d. Quasi-expérimentale
e. Cross-over

**15** **En résumé, vous diriez de l'étude qu'elle**

a. Présente un manque évident de puissance statistique
b. Justifie de rendre obligatoire l'analyse systématique de cas d'hémorragies sévères du post-partum dans toutes les maternités françaises
c. Aurait pu bénéficier de l'ajout d'une mesure du critère principal dans les deux bras de l'essai avant la mise en place de l'intervention
d. Permet d'objectiver que l'intervention n'a pas eu d'effet
e. Présente un niveau de preuve plus élevé qu'une étude avant-après



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 1**

**Réponse exacte : D**

*Commentaires* L'hypothèse principale de l'étude est explicitement présentée à la fin de la section « Introduction ».

**Question 2**

**Réponses exactes : A – D – E**

**Inadmissible : C**

*Commentaires* Il s'agissait d'un essai randomisé en clusters en deux bras parallèles comparant deux groupes de maternités tirées au sort pour bénéficier ou non de l'intervention, sans réalisation préalable d'une phase où les critères de jugements sont évalués avant la mise en place de l'intervention. Les patientes n'étaient donc pas tirées au sort individuellement pour appartenir à l'un ou l'autre des bras de l'essai, mais collectivement en fonction de leur maternité de prise en charge.

**Question 3**

**Réponses exactes : B – E**

*Commentaires* Le cluster ou « grappe » correspondait à l'unité de regroupement des patientes, à savoir leur maternité de prise en charge, quel que soit le groupe (ou bras) dans lequel la maternité était tirée au sort et son réseau d'appartenance. L'intervention était donc randomisée au niveau des maternités et non des patientes, chaque maternité (et l'ensemble des patientes y étant prises en charge) appartenant soit au bras intervention, soit au bras contrôle de l'étude. Un tel schéma d'étude devait éviter le risque de contamination entre les bras de l'étude, à savoir que les patientes prises en charge dans les maternités du bras contrôle bénéficient de l'intervention après la randomisation. En effet, si la randomisation avait été conduite au niveau individuel de chaque patiente, une même maternité aurait pris en charge à la fois des patientes bénéficiant de l'intervention et des patientes n'en bénéficiant pas, ce qui n'était pas faisable compte tenu de la nature de l'intervention ciblant les pratiques cliniques d'une maternité dans son intégralité.

**Question 4**

**Réponses exactes : A – C – D**

**Inadmissible : B**

*Commentaires* L'intervention n'a été mise en place que dans les maternités tirées au sort dans le bras intervention de l'étude. Elle ciblait directement les maternités (et non les patientes) afin de permettre des modifications de pratiques pouvant améliorer indirectement le résultat de la prise en charge des patientes.

**Question 5**

**Réponses exactes : A – C – D**

*Commentaires* Le critère de jugement principal mesurait la survenue d'hémorragies sévères du post-partum à partir d'un retour manuel au dossier médical des patientes par des assistants de recherche clinique. Il s'agissait d'un taux exprimé sous la forme d'un rapport entre le nombre d'accouchements avec survenue d'une hémorragie sévère du post-partum et le nombre total d'accouchements réalisés pendant la période d'étude. La survenue d'une hémorragie sévère du post-partum était mesurée en combinant divers évènements pouvant survenir suite à la perte



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

sanguine (transfusion, embolisation artérielle, hystérectomie, etc.). Il s'agissait d'un critère de résultat clinique et non d'un critère de conformité des pratiques à un référentiel (critère de jugement secondaire de l'étude).

### Question 6

**Réponses exactes : B – E**
**Inadmissible : D**

*Commentaires* La randomisation était stratifiée selon les caractéristiques des maternités (leur réseau périnatal d'appartenance et leur volume d'accouchement) et non des patientes. Par contre, elle n'objectivait pas la comparabilité initiale du critère de jugement principal entre les deux bras de l'essai car cette information n'était pas disponible (absence de phase pré-intervention mesurant initialement la valeur du critère de jugement dans l'ensemble des maternités) pour guider la randomisation. La randomisation était donc réalisée au niveau des maternités, et non des patientes qui appartenaient au bras intervention ou contrôle de l'essai en fonction de leur maternité de prise en charge. Par conséquent, la randomisation a été réalisée en une fois afin de décider de la mise en place ou non de l'intervention dans les maternités, et pas de manière répétée au moment de l'inclusion de chaque nouvelle patiente dans l'essai.

### Question 7

**Réponses exactes : A – B – E**

*Commentaires* Le calcul de la taille d'échantillon visait à objectiver une différence absolue de 0,4 point sur le critère principal entre les deux bras de l'essai en faisant l'hypothèse que 1 340 femmes seraient incluses en moyenne dans chacune des maternités.

### Question 8

**Réponses exactes : A – C**
**Inadmissible : D**

*Commentaires* La population finalement analysée comprenait des maternités avec des volumes annuels d'accouchement très variables allant de 140 à 4 100 (*cf.* Figure 1. Flow diagram) et reposait sur une approche de type « intention de traiter ». Cette population était représentative des accouchements réalisés dans les maternités des 6 réseaux de périnatalité participant à l'essai. Pour être représentative des accouchements en France, il aurait fallu constituer un échantillon représentatif des réseaux de France, par tirage au sort par exemple.

### Question 9

**Réponse exacte : C**
**Inadmissible : B**

*Commentaires* L'interprétation des résultats ne permettait pas de savoir si les deux bras de l'étude étaient comparables initialement sur le critère principal en l'absence de phase pré-intervention. Par ailleurs, en l'absence de différence, elle ne permettait pas de conclure à l'efficacité de l'intervention sur le critère principal. Par contre, certaines pratiques recommandées étaient davantage appliquées dans le bras intervention, tel l'appel à l'aide d'un collègue senior.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 10**

**Réponses exactes : B – D**

*Commentaires* L'appel à l'aide d'un collègue senior était 1,65 fois plus fréquent dans le bras intervention comparativement au bras contrôle de l'essai.

**Question 11**

**Réponse exacte : D**

**Inadmissible : B**

*Commentaires* Il n'existait pas de différence de tendance significative entre les deux groupes sur le critère principal, ni d'amélioration statistiquement significative dans le bras intervention de l'étude. En l'absence de phase pré-intervention, la tendance préexistante à la mise en place de l'intervention était inconnue. Par ailleurs, les raisons d'une telle absence de différence entre les bras demeuraient inconnues, notamment la survenue éventuelle d'un biais de contamination.

**Question 12**

**Réponses exactes : A – B – C – D**

*Commentaires* La conduite d'un essai en clusters vise à limiter le risque de contamination entre les bras. Cependant, il est probable qu'en raison de l'impossibilité de mettre en œuvre l'intervention en insu des professionnels, certaines maternités du bras contrôle aient eu des informations sur l'intervention mise en place et aient cherché à la copier. Par ailleurs, en l'absence de phase pré-intervention, nous ne pouvons affirmer que la valeur du critère de jugement principal était similaire entre les bras avant la mise en place de l'intervention. Il est donc possible que le taux d'hémorragies sévères du post-partum était plus élevé au départ dans les maternités du bras intervention comparativement à celles du bras contrôle, rendant plus difficile la mise en évidence d'un effet de l'intervention.

**Question 13**

**Réponses exactes : A – B – C**

*Commentaires* L'essai a été financé par le Ministère de la Santé français dans le cadre du Programme Hospitalier de Recherche Clinique et publié dans une revue de gynécologie-obstétrique. Le groupe des co-investigateurs avec leurs affiliations respectives apparaît en page 1278 du papier, tandis que les aspects éthiques (Details of ethics approval) et la source de financement (funding) sont précisés en pages 1285-1286.

**Question 14**

**Réponses exactes : B – C**

*Commentaires* Il s'agissait d'une étude expérimentale randomisée en clusters sans insu.

**Question 15**

**Réponses exactes : C – E**

**Inadmissible : B**

*Commentaires* En l'absence d'effet démontré de l'intervention, cet essai ne justifie pas la mise en place d'une analyse systématique de cas d'hémorragies sévères du post-partum dans toutes les maternités françaises. Néanmoins, l'étude ne permet pas de conclure à l'absence d'effet de l'intervention, un déséquilibre initial des bras sur le critère principal ou un biais de contamina-



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

tion ayant pu survenir et expliquer l'absence de différence. Enfin, l'essai randomisé représente aujourd'hui le schéma d'étude avec le niveau de preuve le plus élevé et est clairement supérieur à celui apporté par une étude avant-après.

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 13 Essai d'intervention

Étienne Audureau

| | |
|---|---|
| **Titre de l'article** | **A Cluster-Randomized Trial to Reduce Cesarean Delivery Rates in Quebec** |
| **Auteurs** | Chaillet N, Dumont A, Abrahamowicz M, Pasquier JC, Audibert F, Monnier P, Abenhaim HA, Dubé E, Dugas M, Burne R, Fraser WD ; QUARISMA Trial Research Group |
| **Référence** | N Engl J Med. 2015 ; 372(18):1710-21. |
| **Lien vers l'article** | http://www.nejm.org/doi/10.1056/NEJMoa1407120 |

**1** **L'étude effectuée est caractérisée par le(s) mot(s) clé(s) suivant(s)**

a. Randomisée
b. Expérimentale
c. Quasi-expérimentale
d. De supériorité
e. Observationnelle

**2** **L'objectif principal de l'étude était de**

a. Décrire les pratiques d'accouchement par césarienne au Québec
b. Identifier les facteurs de risque de pratiques inadaptées en milieu obstétrique au Québec
c. Démontrer l'efficacité d'une intervention multifacette sur la diminution du taux de césarienne
d. Démontrer l'efficacité d'une intervention multifacette pour diminuer l'incidence des complications liées aux accouchements par césarienne
e. Démontrer la valeur pronostique péjorative des accouchements par césarienne

**3** **Pour répondre à l'objectif principal de démonstration de l'efficacité d'une intervention multifacette, les auteurs auraient également pu conduire**

a. Une étude exposés/non-exposés prospective
b. Une étude de type « avant-après » avec groupe contrôle sans intervention
c. Un essai contrôlé randomisé avec randomisation individuelle des patientes
d. Une étude de série chronologique basée sur l'évolution des pratiques à long terme
e. Aucune réponse n'est exacte



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **À propos de la randomisation, il est vrai que**

a. Elle était stratifiée sur le type de catégorie professionnelle assurant la prise en charge (sage-femme/obstétricien)
b. Elle était stratifiée sur le type de centre participant (communautaire, régional, tertiaire)
c. La randomisation par blocs a permis d'équilibrer le nombre de centres inclus dans chacun des deux groupes
d. La randomisation par blocs a permis de diminuer le biais de contamination entre patientes de groupe différent
e. Aurait dû également tenir compte des facteurs pronostiques connus chez les patientes participantes à l'étude

**5** **Les critères d'éligibilité de la population d'étude**

a. Comprenaient pour les centres un nombre annuel minimal d'accouchements
b. Comprenaient pour les centres un taux maximal d'accouchement par césarienne de 17 %
c. Comprenaient pour les patientes l'absence de grossesse multiple
d. Comprenaient pour les patientes l'absence de prématurité inférieure à 24 semaines de grossesse
e. Ont permis aux investigateurs d'inclure tout type de maternité dans la province du Québec, Canada

**6** **Le nombre de sujets nécessaire présenté par les auteurs**

a. Est calculé avec une puissance et un risque de première espèce de 20 %
b. S'appuie sur un coefficient de corrélation intra-classe égal au taux moyen d'accouchement par césarienne au Québec en 2006-2007
c. Tient compte de la structure hiérarchique des données (patients nichés au sein des centres)
d. S'appuie sur la comparaison de l'évolution du taux de césarienne entre les périodes pré- et post-intervention
e. S'appuie sur la comparaison de la valeur finale du taux de césarienne sur la période post-intervention entre les deux groupes

**7** **L'intervention multifacette combinait**

a. Un recours à un *leader* d'opinion local afin d'appuyer les recommandations
b. Une formation initiale sur les bonnes pratiques
c. Des outils informatisés d'aide à la décision
d. Des aides financières incitatives pour l'implémentation des bonnes pratiques
e. Des audits cliniques répétés avec retour d'information vers les soignants



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**8** **Le recueil des informations pour les critères de jugement**

a. Était réalisé à partir d'auto-questionnaires de santé complétés par les patientes
b. Était réalisé à partir des dossiers médicaux en insu du groupe de randomisation
c. Était réalisé de manière identique pour les deux groupes randomisés et les périodes pré- et post-intervention
d. N'exposait pas au biais de mémorisation
e. Exposait à un biais d'attrition

**9** **La méthode d'analyse statistique du critère de jugement principal**

a. Était basée sur une approche en *generalized-estimating-equations* tenant compte de la structure emboîtée des données
b. Aurait pu s'appuyer sur une analyse en modèle de Cox à risques proportionnels
c. Prend en compte la corrélation entre les patients d'un même hôpital
d. Aurait pu s'appuyer sur une comparaison de pourcentages par test du $Chi^2$
e. Aurait pu s'appuyer sur une régression linéaire par la méthode des moindres carrés

**10** **L'Odds ratio (OR) ajusté se rapportant à l'analyse du critère de jugement principal présenté dans le tableau 2**

a. Est statistiquement significatif en faveur du groupe intervention car OR = -1,8 ($IC_{95\,\%}$ -3,8 ;-0,2)
b. Est statistiquement significatif en faveur du groupe intervention car OR = 0,90 ($IC_{95\,\%}$ 0,80 ; 0,99)
c. Est associé à une diminution de 0,9 % du taux de césarienne pour le groupe « intervention » entre la phase pré- et post-intervention
d. Correspond à une réduction relative de 10 % du taux de césarienne dans le groupe intervention par rapport au groupe contrôle
e. Est basé sur l'interaction statistique entre le temps d'analyse (pré-/post-intervention) et le groupe randomisé (contrôle/intervention)

**11** **Les résultats stratifiés selon le risque de la grossesse présentés dans le tableau 2**

a. Permettent de confirmer le résultat significatif de la question 10 chez les patientes ayant une grossesse à bas niveau de risque
b. Permettent de confirmer le résultat significatif de la question 10 chez les patientes ayant une grossesse à haut niveau de risque
c. Montrent une diminution statistiquement plus importante dans le groupe contrôle chez les patientes ayant une grossesse à haut niveau de risque
d. Correspondent à une analyse post-hoc portant sur des sous-groupes non pré-spécifiés
e. Permettent de poser une conclusion clinique pour le sous-groupe où une différence significative est retrouvée



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**12** **L'analyse en per protocole**

a. Portait sur les données après imputation des données manquantes
b. Excluait les patientes avec grossesse multiple
c. Excluait les patientes perdues de vue ou sorties d'étude
d. Excluait les centres n'ayant pas implémenté l'intervention
e. Confirmait le résultat obtenu en ITT pour le critère de jugement principal avec un OR = 0,82 (0,72 ; 0,89)

**13** **L'analyse principale était réalisée après ajustement**

a. Sur des facteurs de risque individuels associés à la pratique d'une césarienne
b. Sur des facteurs de risque de niveau centre associés à la pratique d'une césarienne
c. Afin de contrôler les biais de classement éventuels
d. Afin de contrôler les biais de confusion éventuels
e. Afin de tenir compte du devis de l'étude

**14** **Le risque de complications mineures et majeures**

a. Était significativement diminué dans le groupe intervention comparativement au groupe contrôle pour les complications maternelles
b. Était significativement diminué dans le groupe intervention comparativement au groupe contrôle pour les complications néonatales
c. Était augmenté de 2 % dans le groupe intervention pour les complications néonatales mineures
d. Constituait des critères de jugement secondaires pré-spécifiés
e. Constituait des critères de stratification pré-spécifiés

**15** **D'après les auteurs :**

a. La possibilité de généraliser les résultats est renforcée par la simplicité relative de la mise en œuvre de l'intervention
b. La possibilité de généralisation des résultats est limitée par la non-inclusion des centres privés
c. L'efficacité de chacune des composantes de l'intervention n'a pas pu être évaluée séparément
d. L'intervention a permis d'obtenir une diminution statistiquement significative et cliniquement importante du taux de césarienne
e. L'intervention a permis d'obtenir une diminution du taux de césarienne sans augmentation de la morbidité et mortalité.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 1

**Réponses exactes : A – B – D**

*Commentaires* Il s'agit d'une étude expérimentale randomisée multicentrique en deux bras parallèles visant à démontrer l'efficacité (supériorité) d'une intervention comparativement à un groupe contrôle. Une randomisation a été effectuée pour assurer une allocation aléatoire de l'intervention entre les groupes, définissant donc le contexte expérimental. Les études quasi-expérimentales – également souvent conduites pour évaluer des interventions d'amélioration des pratiques – concernent des études non randomisées, par exemple de type « avant-après » ou « ici-ailleurs », parfois avec groupe contrôle mais sans randomisation et donc avec un niveau de preuve inférieur.

### Question 2

**Réponse exacte : C**

*Commentaires* S'agissant d'une étude randomisée, l'objectif principal précisé comme attendu par les auteurs à la fin de l'introduction est de démontrer l'efficacité de l'intervention sur la diminution du taux d'accouchement par césarienne. Les propositions alternatives A-B-E se rapportent à des objectifs d'épidémiologie descriptive ou analytique. La réponse D correspond aux objectifs secondaires et critères de jugement secondaires associés et non à l'objectif principal de modification des pratiques. Le mot-clé « multifacette » s'applique à l'intervention menée qui comporte plusieurs composantes (facettes).

### Question 3

**Réponse exacte : E**

*Commentaires* Les propositions A-B-D concernent des types d'études non randomisées ne permettant pas de répondre à l'objectif principal de démonstration d'efficacité.

Un essai randomisé avec randomisation individuelle des patientes serait ici inadapté, du fait des modalités de l'intervention comprenant notamment des audits avec retour d'information aux équipes soignantes, empêchant de fait l'analyse de l'efficacité patiente par patiente sans biais de contamination majeur (*cf.* question 4). Une randomisation par centres est ici indispensable.

### Question 4

**Réponses exactes : B – C**

*Commentaires* La stratification était réalisée sur le type de centre participant, facteur potentiellement associé aux pratiques d'accouchement, afin de garantir l'équilibre entre groupes randomisés sur ce critère. La randomisation en blocs assure quant à elle un équilibre sur le nombre de centres par strate et groupe. A titre d'exemple, l'utilisation de blocs de 4 permet d'obtenir pour une strate donnée [ex : hôpitaux tertiaires] une succession d'allocations aléatoires comportant à chaque fois 2 centres « contrôle » et 2 centres « intervention » (ex : Contrôle – Intervention – Intervention – Contrôle ; Intervention-intervention-contrôle-contrôle ; etc.), jusqu'à ce que tous les centres de la strate soient randomisés.

Des facteurs individuels pronostiques importants liés aux caractéristiques des patientes participantes pourraient également influencer les taux d'accouchement par césarienne. Il n'est cependant pas envisageable de prendre en compte ces facteurs dans la stratégie de randomisation, puisqu'il s'agit ici d'une randomisation par hôpital (en cluster ou grappe), et non individuelle, les patientes étant incluses dans un second temps après que les centres étaient randomisés dans l'un des deux groupes expérimentaux. Certaines études réalisent un recueil rétrospectif d'informations en amont de l'étude afin de décrire les caractéristiques des



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

patientes « typiquement » prises en charge dans le centre ; ces informations peuvent alors être utilisées pour la stratification des centres mais ne concernent de toute façon pas les patientes participant réellement à l'étude.

Attention, ne pas confondre la notion de bloc, qui vise à équilibrer l'effectif des groupes de randomisation, et la notion de randomisation en cluster ou grappes qui s'oppose à la randomisation individuelle. Une grappe est un ici un hôpital.

La stratification de la randomisation n'a pas de lien avec le biais de contamination. Le biais de contamination s'applique aux études visant à étudier l'effet d'interventions ou traitements sur la base d'une randomisation individuelle des participants, alors qu'il existe un risque de contamination inter-groupes du fait par exemple de la difficulté pour les médecins prenant en charge des patients des deux groupes de ne pas appliquer les recommandations de prise en charge ou la nouvelle intervention aux patients du groupe contrôle, ou lorsque l'on étudie des agents infectieux contagieux où la contamination inter-groupes devient alors une réalité médicale rendant difficile l'évaluation de l'effet propre du traitement ou de l'intervention.

### Question 5

**Réponses exactes : A – D**

*Commentaires* Les critères d'éligibilité des centres et patientes sont décrits dans le paragraphe « Hospitals and participants » au début de la section « Méthodes ». La proposition B est inversée (bonne réponse « taux minimal d'accouchement »). La proposition E est fausse, l'étude n'incluant que des centres publics avec une activité minimale de 300 accouchements annuels.

### Question 6

**Réponses exactes : C – D**

*Commentaires* Le calcul du nombre de sujets nécessaire pour un essai randomisé en clusters doit tenir compte de la structure hiérarchique des données, où les participants sont « emboîtés » au sein d'une unité hiérarchique supérieure, par exemple patients pris en charge par un même médecin, patients pris en charge au sein d'un hôpital, enfants scolarisés au sein d'une même école, etc. Il existe alors une certaine corrélation entre les participants issus d'un même cluster, reflétant le degré de ressemblance entre les sujets d'un même cluster. Cette « ressemblance » s'exprime via le coefficient de corrélation intraclasse, dont une valeur élevée signe un niveau de corrélation important entre sujets (min 0 ; max 1). Il est en général faible mais doit cependant être pris en compte lors du calcul du nombre de sujets nécessaire et au moment de l'analyse (*cf.* plus loin). Pour le nombre de sujets, la corrélation représente une nuisance qui nécessite une augmentation de l'effectif requis par rapport à un essai à randomisation individuelle. Si le calcul d'effectif ne tient pas compte de cette corrélation, l'essai aura une puissance statistique insuffisante par rapport à l'objectif initial.

Le risque de première espèce (alpha) est ici de 5 % en test bilatéral.

Le calcul s'appuie sur le critère de jugement principal qui porte sur la comparaison de l'évolution (diminution) du taux d'accouchement par césarienne entre les groupes randomisés. Ce critère ne s'appuie donc pas seulement sur la valeur finale en post-intervention mais tient aussi compte du niveau basal pré-intervention.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 7

**Réponses exactes : A – B – E**

*Commentaires* Les détails concernant l'intervention sont donnés dans le paragraphe dédié « Intervention » de la section « Méthodes ».

### Question 8

**Réponses exactes : C – D**

*Commentaires* Le recueil était basé sur les dossiers médicaux, non sujets à biais de mémorisation, mais les professionnels en charge de ce recueil n'étaient pas en aveugle du groupe de randomisation (avant-dernier paragraphe de la section « Study design »). Le biais d'attrition est un biais de sélection se constituant en cours d'étude du fait des perdus de vue et non un biais de classement.

### Question 9

**Réponses exactes : A – C**

*Commentaires* L'analyse des essais randomisés en clusters doit tenir compte de la structure hiérarchique des données et de la corrélation existant entre sujets issus des mêmes unités (ex : cabinets de médecine générale, école, hôpitaux, etc.). En cas de non prise en compte de cette corrélation, le risque de conclure à tort à une différence entre les groupes de traitement est augmenté. Les méthodes d'analyse statistique permettant de tenir compte de cette corrélation incluent les modèles dits en « generalized estimating equations » (GEE) et les modèles dits à effets « mixtes » (ou modèles « multiniveau », ou « à effets fixes et aléatoires »), pouvant s'appliquer tant pour des critères de jugement binaires comme dans cette étude (extension des modèles de régression logistique, ex : accouchement par césarienne, complication majeure, etc.) ou pour des critères quantitatifs (extension des modèles de régression linéaire, ex : tension artérielle, indice de masse corporelle, etc.).

### Question 10

**Réponses exactes : B – D – E**

*Commentaires* -1,8 correspond à la différence absolue de risque ajustée, obtenue par régression logistique multivariée en GEE et correspondant à la différence entre groupes randomisés pour l'évolution du taux de césarienne entre la phase initiale et la phase post-intervention. L'OR d'intérêt correspond à l'analyse en intention de traiter (ITT) dans la partie haute du tableau ; l'intervalle de confiance ne comprend pas 1 (et le degré de signification p est < 0,05).

L'OR correspond à l'interaction entre la période d'analyse pré- VS post-intervention et le groupe de randomisation contrôle VS intervention. Une interaction significative indique que la diminution du taux de césarienne avant-après intervention est plus importante pour le groupe « intervention » que pour le groupe contrôle (d'où « interaction » période * groupe).

### Question 11

**Réponses exactes : A – D**

*Commentaires* L'OR est de 0,80 (0,65 ; 0,97) avec un p = 0,03 chez les patientes ayant une grossesse à bas niveau de risque, alors qu'aucune différence statistiquement significative n'est retrouvée chez les femmes ayant une grossesse à haut niveau de risque (p = 0,35).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Un test d'interaction est réalisé avant de stratifier sur ce critère, revenant à se demander si la différence contrôle VS intervention varie selon le niveau de risque de la grossesse. Cette analyse en sous-groupes n'est néanmoins pas précisée dans la partie « Méthodes » comme étant prévue dès le départ lors de l'écriture du protocole (« prespecified » ; le protocole détaillé en français et accessible en « Supplementary Material » sur le site du journal confirme cette absence). Il s'agit donc d'une analyse post-hoc décidée après le recueil des données qui ne permet pas de conclusion clinique mais a seulement une valeur exploratoire.

### Question 12

**Réponses exactes : D – E**

*Commentaires* L'analyse en per protocole portait sur les centres ayant effectivement implémenté l'intervention (« adherence » définie par les auteurs dans le dernier paragraphe de la section « Méthodes ») et a donc exclu 4 centres sur les 16 du groupe intervention qui ne remplissait pas les critères définis par les auteurs. Le résultat obtenu en ITT est confirmé et même accentué (comme attendu après exclusion des centres non compliants).

L'imputation des données manquantes concernait l'analyse en ITT, où l'on cherche à conduire l'analyse sur la population randomisée initiale, en limitant au maximum l'impact des données manquantes pour ne pas exclure les patientes sur ce motif.

### Question 13

**Réponses exactes : A – B – D**

*Commentaires* Les facteurs d'ajustement sont spécifiés dans le paragraphe « Statistical analysis » de la section « Méthodes ». Il s'agit de facteurs de confusion que l'analyse multivariée permet de contrôler.

### Question 14

**Réponses exactes : B – C – D**

*Commentaires* Ces résultats sont détaillés dans le tableau 4.

### Question 15

**Réponses exactes : A – C – E**

*Commentaires* Même si la population couvre un très grand nombre d'accouchements pris en charge au Québec (65 % d'après les auteurs), la non-inclusion des centres privés constitue un biais de sélection non discuté par les auteurs. Pourtant, on pourrait imaginer que l'efficacité de l'intervention puisse varier selon le statut public-privé de l'hôpital, mais également des attentes des patientes et du contexte culturel (ex : taux très élevés des césariennes au Brésil), facteurs pouvant limiter la validité externe des résultats de l'étude. Pour la réponse D, la différence reste cliniquement faible malgré la différence statistique retrouvée sans doute plus facilement grâce aux effectifs importants recrutés et à la puissance statistique résultante élevée.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 14 Essai de prévention

Jérémie Jegu

| | |
|---|---|
| **Titre de l'article** | **Tamoxifen for prevention of breast cancer : extended long-termfollow-up of the IBIS-I breast cancer prevention trial** |
| **Auteurs** | Cuzick J., Sestak I., Cawthorn S., Hamed H., Holli K., Howell A., Forbes J., on behalf of the IBIS-I Investigators |
| **Référence** | Lancet Oncology. janv 2015 ; 16(1):67-75 |
| **Lien vers l'article** | **http://www.thelancet.com/journals/lanonc/article/PIIS1470-2045(14)71171-4/abstract** |

**1** **Les investigateurs de cette étude ont choisi d'inclure des femmes à haut risque de cancer du sein. Pour quelle(s) raison(s) ?**

a. Pour améliorer la précision des estimations obtenues dans le cadre de l'étude
b. Parce que ces femmes sont représentatives de l'ensemble des femmes de la population générale
c. Pour observer davantage d'événements lors du suivi
d. Pour avoir plus de chance de mettre en évidence un effet du traitement si celui-ci existe réellement
e. Parce que les études précédemment publiées sur ce sujet portaient sur ce type de population

**2** **Cette étude comporte l'analyse des effets à long terme (> 10 ans) de la prise de Tamoxifen en prévention du cancer du sein. Pour quelle(s) raison(s) ?**

a. Parce qu'un effet protecteur sur le risque de cancer du sein a déjà été démontré lors des 10 premières années de suivi
b. Parce qu'un effet protecteur sur la mortalité a déjà été démontré lors des 10 premières années de suivi
c. Pour évaluer la persistance d'un effet protecteur sur le risque de cancer du sein à long terme
d. Pour disposer de données sur la mortalité à long terme
e. Pour évaluer les effets indésirables à long terme

**3** **Parmi les qualificatifs suivants, lequel (lesquels) est (sont) pertinent(s) pour cette étude ?**

a. Cohorte étiologique
b. Randomisation stratifiée
c. Rétrospectif
d. Simple aveugle
e. Double aveugle



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **L'allocation du traitement a été effectuée par tirage au sort dans cette étude. Pour quelle(s) raison(s) ?**

a. Rendre la mesure du critère de jugement principal similaire entre les groupes
b. Assurer la représentativité des conclusions de l'étude par rapport à la population cible
c. Rendre parfaitement comparables les deux groupes
d. Limiter autant que possible les déséquilibres de caractéristiques entre les deux groupes
e. Renforcer la puissance statistique de l'étude

**5** **Le critère de jugement principal est ici la survenue d'un cancer du sein invasif ou d'un carcinome in situ canalaire (ductal carcinoma in situ). Parmi les propositions suivantes, laquelle (lesquelles) est (sont) correcte(s) ?**

a. Il s'agit d'un critère de substitution par rapport à la mortalité
b. Il est lié à la mortalité puisque la présence d'un cancer du sein augmente le risque de décès
c. C'est un critère composite
d. Ce type de critère permet d'améliorer la puissance de l'étude
e. Ce type de critère permet de travailler sur une population de taille moins importante et pour une durée de suivi plus courte, ce qui est important dans le cadre d'un essai de prévention

**6** **La séquence d'allocation aléatoire du traitement a été préparée avant le début de l'étude IBIS-S. Pour quelle(s) raison(s) ?**

a. Afin de garantir le triple aveugle
b. Pour préserver le caractère totalement aléatoire de la génération de cette séquence d'allocation
c. Pour faciliter le déroulement pratique de l'étude
d. Pour garantir l'obtention de groupes de patients ayant des caractéristiques cliniques équilibrées
e. Pour obtenir une séquence d'allocation identique dans l'ensemble des centres

**7** **La mortalité toutes causes a été étudiée en tant que critère secondaire ici. Pour quelle(s) raison(s) ?**

a. Le décès est un événement qui survient à trop long terme
b. Ce n'est pas un critère important du point de vue clinique
c. C'est un critère difficile à mesurer
d. C'est un critère trop peu fréquent pour disposer d'une puissance statistique suffisante
e. C'est un critère important compte tenu du critère principal choisi

**8** **Au début de la section résultats, les auteurs signalent que 15 femmes ont été finalement non incluses car elles présentaient un diagnostic de cancer du sein antérieur. Pour quelle(s) raison(s) ?**

a. Il faut que les participantes soient indemnes de l'événement étudié en début de suivi
b. Il s'agit d'un critère d'exclusion
c. La survenue d'un cancer du sein est précisément ce que l'on cherche à prévenir dans cet essai
d. Les femmes avec des antécédents de cancer ne font pas partie de la population cible
e. L'intervention serait trop tardive pour ces patientes



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**9** **Parmi les propositions suivantes, laquelle (lesquelles) est- (sont-) correcte(s) concernant la figure 1 ?**

a. Ces courbes ont été construites en utilisant la méthode de Kaplan-Meier
b. Les courbes présentées sont ajustées sur différents facteurs de confusion
c. Il s'agit de courbes montrant la survenue de l'événement au cours du temps
d. Les deux lignes de chiffres indiquées en bas de la figure correspondent au nombre d'événements au cours du temps
e. Le chiffre 12,3 % indiqué tout en haut à droite signifie que 12,3 % des participantes sont décédées au cours de 20 ans de suivi dans le groupe placebo

**10** **Dans le tableau 1, quelle(s) proposition(s) est (sont)-elle(s) correcte(s) concernant le nombre total de cancers du sein (2e ligne) ?**

a. Pour l'ensemble du suivi, le rapport de cotes est de 0,71 avec un intervalle de confiance à 95 % (0,60-0,83)
b. Pour l'ensemble du suivi, le risque de survenue d'un cancer du sein est diminué de 29 % dans le groupe Tamoxifen
c. Le groupe placebo présente un risque plus important de développer un cancer du sein par rapport au groupe Tamoxifen
d. L'effet du Tamoxifen est statistiquement significatif lors des 10 premières années de suivi
e. L'effet du Tamoxifen est statistiquement significatif après 10 ans de suivi

**11** **Parmi les propositions suivantes laquelle (lesquelles) est (sont) correcte(s) concernant le tableau 1 ?**

a. On n'observe pas d'effet protecteur du Tamoxifen sur le risque de cancers du sein négatifs aux récepteurs aux œstrogènes (oestrogen receptor-negative cancer)
b. L'effet protecteur du Tamoxifen est statistiquement significatif pour les carcinomes canalaires in situ quelle que soit la durée de suivi
c. L'effet protecteur du Tamoxifen est plus marqué pour les cancers du sein positifs aux récepteurs aux œstrogènes
d. Il est possible que le Tamoxifen entraîne une augmentation du risque de cancers du sein négatifs aux récepteurs aux œstrogènes pour les durées de suivi supérieures à 10 ans, mais cet effet n'est pas statistiquement significatif
e. Il est possible que le Tamoxifen entraîne une augmentation du risque de cancer du sein de haut grade pour les durées de suivi supérieures à 10 ans, mais cet effet n'est pas statistiquement significatif



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**12** **Indiquez parmi les propositions suivantes laquelle convient aux résultats de cette étude**

a. Le nombre nécessaire de patients à dépister pendant une durée de 20 ans est de 22 pour détecter 1 cancer du sein
b. Le nombre nécessaire de patients à dépister pendant une durée de 5 ans est de 22 pour détecter 1 cancer du sein
c. Tous les patientes traitées par Tamoxifen seront totalement protégées du risque de cancer du sein
d. Il faudrait traiter 22 patientes par Tamoxifen sur une durée de 20 ans pour prévenir 1 cancer du sein
e. Il faudrait traiter 22 patientes par Tamoxifen sur une durée de 5 ans pour prévenir 1 cancer du sein

**13** **Parmi les propositions suivantes, laquelle (lesquelles) est (sont) correcte(s) concernant le tableau 3 ?**

a. Le nombre de décès lors du suivi était supérieur dans le groupe traité par Tamoxifen
b. On n'observe pas de différence notable du nombre de décès entre les groupes pour les deux périodes de suivi
c. La proportion de décès par cancer du poumon était moins importante dans le groupe Tamoxifen
d. La proportion de décès par cancer du sein était moins importante dans le groupe Tamoxifen
e. Le nombre de décès par cancer du sein était deux fois plus élevé dans le groupe Tamoxifen par rapport au groupe placebo pour les durées de suivi supérieures à 10 ans

**14** **Parmi les propositions suivantes, laquelle (lesquelles) est (sont) évoquée(s) par les auteurs de cet article à propos de la mortalité ?**

a. L'excès de mortalité en cas de traitement par Tamoxifen avait déjà été rapporté dans des études précédentes
b. L'excès de mortalité par cancer de l'endomètre en cas de traitement par Tamoxifen avait déjà été rapporté dans des études précédentes
c. Le nombre de décès était identique dans les deux groupes après 10 ans de suivi
d. 5 décès par cancer de l'endomètre ont été constatés dans le groupe Tamoxifen contre 0 dans le groupe placebo
e. L'absence de démonstration d'un effet protecteur significatif du Tamoxifen sur la mortalité est due à un manque de puissance statistique

**15** **Une femme âgée de 60 ans sans antécédent familial de cancer du sein et sans antécédent personnel de lésion bénigne du sein vous demande lors d'une consultation s'il serait intéressant qu'elle prenne du Tamoxifen pour diminuer son risque de cancer du sein. Quelle réponse principale lui donnez-vous en vous basant uniquement sur les résultats de cette étude ?**

a. Vous préférez ne pas vous prononcer
b. Que la prise de Tamoxifen peut effectivement réduire son risque de cancer du sein
c. Que le rapport bénéfice/risque n'est pas en faveur du Tamoxifen
d. Que le Tamoxifen peut réduire son risque de décès par cancer du sein
e. Que le Tamoxifen peut réduire son risque de décès par cancer du poumon



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 1

**Réponses exactes : C – D – E**

**Inadmissible : B**

*Commentaires* Le fait d'inclure une population à haut risque dans le cadre d'un essai de prévention permet d'augmenter le nombre d'événements observés lors du suivi (réponse C), ici la survenue d'un cancer ou d'une lésion in situ du sein, ce qui améliore la puissance de l'étude (capacité à rejeter l'hypothèse nulle si celle-ci est fausse – réponse D). Dans le cas particulier de cette étude, les essais cliniques précédemment publiés sur l'intérêt du Tamoxifen en prévention primaire portaient effectivement sur des populations de femmes à haut risque de cancer du sein (réponse E). Le fait d'étudier une population à haut risque n'améliorera pas la précision des estimations obtenues (réponse A). En effet, la précision des estimations dépend du nombre de participantes incluses et non du nombre d'événements observés. Les femmes à haut risque de cancer du sein ne peuvent en aucun cas être considérées comme représentatives de l'ensemble des femmes de la population générale (réponse B).

### Question 2

**Réponses exactes : A – C – D – E**

*Commentaires* Les essais randomisés précédemment publiés ont déjà mis en évidence un effet protecteur à court terme (<10 ans) de la prise de Tamoxifen sur le risque de cancer du sein dans cette population (réponse A). L'originalité de l'article porte ici principalement sur l'étude de la persistance à long terme de cet effet protecteur sur le risque de cancer du sein (réponse C). L'étude des effets indésirables à long terme du Tamoxifen et de la mortalité font partie des critères secondaires de cette étude (réponse D et E). Il n'y a pas d'effet protecteur démontré sur la mortalité dans les 10 premières années (réponse B).

### Question 3

**Réponses exactes : B – E**

*Commentaires* Il s'agit bien d'un essai de prévention multicentrique randomisé. La randomisation a été stratifiée selon le centre (réponse B) et l'allocation dans le groupe Tamoxifen ou placebo était inconnue de la participante et de l'investigateur (double aveugle – réponse E). Il ne s'agissait pas ici d'un triple aveugle, car le statisticien avait accès aux données concernant le type de traitement alloué. On ne peut pas parler ici d'une étude de cohorte étiologique au sens épidémiologique du terme (ensemble de personnes suivis au cours du temps chez qui l'on étudie l'association entre une exposition préexistante et la survenue d'une maladie), puisque dans le cadre d'un essai clinique l'exposition est contrôlée par l'expérimentateur (réponse A). Pour cette même raison, un essai clinique ne peut pas être rétrospectif (le début de l'étude précède forcément l'allocation des patients dans l'un ou l'autre des groupes, le suivi des patientes étant réalisé dans un second temps, réponse C).

### Question 4

**Réponse exacte : D**

*Commentaires* Un tirage au sort (randomisation) du type de traitement pris par les participantes permet de disposer de groupes dont les caractéristiques sont les plus similaires que possible (réponse D). Si le tirage au sort est réussi, les facteurs de confusion (connus ou non connus) doivent être équilibrés entre les groupes. Une différence d'effet observé peut alors être attribuée à l'intervention, à savoir ici la prise de Tamoxifen. Une procédure de randomisation n'assure pas un équilibre parfait entre les groupes. Malgré une randomisation bien faite, des déséquilibres peuvent



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

persister et rendre les groupes non parfaitement comparables (réponse C). Pour cette raison, les caractéristiques des participants par groupes sont souvent présentées dans un tableau (ici en haut du tableau 1). Des modèles d'analyse multivariés peuvent être utilisés pour limiter l'impact de déséquilibres éventuels. Un tirage au sort du type de traitement n'impacte pas directement la qualité de la mesure des critères de jugement (réponse A). Ce sont les procédures d'aveugle (simple, double, triple) qui permettent de limiter les biais de mesure (ou biais d'information). Il ne faut pas confondre le tirage au sort du type d'intervention (Tamoxifen *versus* placebo ici) et le tirage au sort que l'on pourrait effectuer dans le cadre d'une procédure d'échantillonnage. En effet, lorsqu'il est impossible de travailler avec l'ensemble d'une population donnée, on peut être amené à constituer un échantillon par tirage au sort. Ce tirage au sort a pour but de constituer un échantillon aussi représentatif que possible de la population cible afin de limiter le risque de biais de sélection. Dans le cadre de la présente étude, le tirage au sort vise à limiter les biais de confusion, mais n'impacte pas le risque de biais de sélection et n'assure donc pas la représentativité des résultats à la population cible (réponse B). Enfin, une procédure de randomisation ne permet pas d'augmenter la puissance statistique d'une étude (réponse E). Seule l'observation d'un nombre plus important d'événements lors du suivi pourrait ici augmenter la puissance statistique (effectifs plus importants, population à plus haut risque).

### Question 5

**Réponses exactes : A – C – D – E**

*Commentaires* Le critère de jugement présenté dans le cadre de cette étude est un critère composite (réponse C) car il rassemble des événements de nature différente : des carcinomes in situ (lésions de meilleur pronostic) et des cancers infiltrants. Avec ce type de critère, le nombre total d'événements observés lors du suivi sera plus important, ce qui va améliorer la puissance statistique de l'étude (réponse D). L'interprétation des résultats d'un critère composite doit cependant être prudente : une analyse par sous-type d'événement doit également être présentée (ce qui est le cas dans cet article). En effet, une éventuelle diminution globale du nombre d'événements dans le groupe traitement actif ne devrait pas uniquement concerner les sous-composantes les moins importantes des critères composites (ici les carcinomes in situ). On peut également considérer que c'est un critère de substitution de la mortalité (réponse A). Utiliser un tel critère permet de raccourcir la taille de la population étudiée et la durée globale de suivi, ce qui est très appréciable pour réduire les coûts déjà importants d'un essai de prévention (réponse E). Attention, un critère de substitution n'est pas à coup sûr le reflet de l'événement que l'on cherche à prévenir à plus long terme (réponse B). De manière assez étonnante dans cet article, la prise de Tamoxifen, bien qu'associée à une diminution du critère de jugement principal est associée à une augmentation de la mortalité, ce qui pose un réel problème.

### Question 6

**Réponses exactes : B – C**

*Commentaires* La séquence d'allocation aléatoire a été préparée avant l'étude pour en garantir le caractère totalement aléatoire : aucun événement survenant en cours d'étude ne peut ainsi impacter la génération de cette séquence (réponse B). Disposer d'une liste de randomisation déjà prête est pratique pour le déroulement de l'étude, mais c'est un argument mineur pour la réalisation de ce type de procédure (réponse C). Il n'y a pas de triple aveugle dans cette étude puisque le statisticien était informé du type de traitement alloué (réponse A). Malgré tous les soins apportés à la réalisation d'un tirage au sort, des déséquilibres entre les caractéristiques cliniques des participantes dans chacun des groupes peuvent survenir du simple fait du hasard (réponse D). Comme la randomisation a été stratifiée selon le centre, une séquence d'allocation spécifique par centre a été générée (réponse E).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 7

**Réponse exacte : A – D – E**

**Inadmissible : B**

*Commentaires* La mortalité est, en définitive, ce que l'on cherche à impacter en réduisant l'incidence des cancers du sein dans le cadre de cette étude. C'est cependant un événement qui survient pour des durées de suivi plus importantes (réponse A) et qui est trop peu fréquent pour obtenir une puissance statistique satisfaisante (réponse D). Il est cependant très important d'étudier la mortalité, tout du moins en critère secondaire, pour vérifier que l'impact éventuel du traitement sur le critère principal de jugement s'accompagne d'une réduction du nombre de décès sur le long terme (même non statistiquement significative), ce qui n'est pas le cas dans le cadre de cette étude ! Bien entendu, c'est un critère très important du point de vue clinique (réponse B) et dont la mesure est évidente (réponse C).

### Question 8

**Réponses exactes : A – B – C – D – E**

*Commentaires* Toutes ces propositions sont correctes et expliquent l'exclusion *a posteriori* de ces patientes. En complément, le type d'analyse statistique réalisé, qui est une analyse de survie, n'aurait pas permis l'inclusion de participantes présentant déjà l'événement en début de suivi (durée de suivi nulle). Les patientes présentant un antécédent personnel de lésion bénigne du sein pouvaient être incluses dans cette étude, mais pas celles qui présentaient un antécédent de lésion invasive (cancer) du sein.

### Question 9

**Réponses exactes : A – C**

*Commentaires* Il s'agit bien de courbes construites en utilisant la méthode de Kaplan-Meier (réponse A) et présentant la proportion de participantes présentant l'événement au cours du temps (réponse C). Les courbes de survies construites selon cette méthode ne sont pas ajustées sur les facteurs de confusion potentiels (réponse B). Des analyses multivariées des données censurées peuvent être réalisées avec différents modèles, le plus connu d'entre eux étant le modèle de Cox qui est utilisé dans la suite de cet article. Les deux lignes de chiffres indiquées en bas de la figure rapportent le nombre de participantes aux différents temps de suivi. Il ne pas s'agit pas du nombre d'événements (réponse D) car ces nombres diminuent au fil du temps. Enfin, le chiffre 12,3 % correspond à la proportion de participantes ayant développé un cancer du sein dans les 20 ans de suivi. Le critère présenté ici n'est pas la mortalité.

### Question 10

**Réponses exactes : B – C – D – E**

*Commentaires* Un rapport de risque (Hazard ratio) de 0,71 dans le groupe Tamoxifen par rapport au groupe placebo peut effectivement s'interpréter comme une diminution du risque de survenue de l'événement de 29 % (100 %-71 %) dans le groupe Tamoxifen par rapport au groupe placebo (réponse B). Formulé dans l'autre sens, le groupe placebo présente bien un risque plus important de cancer du sein que le groupe Tamoxifen (réponse C). Cet effet protecteur du Tamoxifen est statistiquement significatif dans les 10 premières années de suivi et après 10 ans de suivi puisque les estimations des rapports de risque sont en dessous de 1 et les intervalles de confiance à 95 % correspondants ne comprennent pas la valeur 1 (réponses D et E).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 11**

**Réponses exactes : A – C – D – E**

*Commentaires* L'effet protecteur du Tamoxifen concerne principalement les cancers du sein positifs aux récepteurs aux œstrogènes (réponse C), et non les cancers du sein négatifs pour ces récepteurs (réponse A). Il est par ailleurs possible que le Tamoxifen augmente sur le long terme (> 10 ans) le risque de cancer de sein positifs au récepteur aux œstrogènes (HR = 2,43) et de cancer du sein de haut grade (HR = 1,27) mais ces effets ne sont pas statistiquement significatifs (réponses C, D et E). L'effet protecteur du Tamoxifen n'est pas statistiquement significatif pour les carcinomes canalaires in situ pour les durées de suivi supérieures à 10 ans : HR 0,91 (0,44 – 1,89) (réponse B).

**Question 12**

**Réponse exacte : E**

*Commentaires* Seule la proposition E est correcte ici, car la période de traitement étudiée dans le cadre de cet essai de prévention est de 5 ans, avec une mesure du critère de jugement principal jusqu'à 20 ans de suivi. Il faudrait traiter 22 patientes (intervalle de confiance à 95 % 19-26) pour espérer prévenir 1 cancer du sein. Il ne s'agit pas d'un dépistage (prévention secondaire des cancers – réponses A et B) et le fait de prendre du Tamoxifen ne diminue pas complètement le risque de développer un cancer du sein lors du suivi (réponse C).

**Question 13**

**Réponses exactes : A – C – E**

*Commentaires* De manière assez étonnante compte tenu de l'analyse du critère de jugement principal qui rapporte une diminution de l'incidence des cancers du sein dans le groupe Tamoxifen, on observe ici un nombre plus important de décès dans le groupe traité par Tamoxifen par rapport au groupe placebo (réponse A). On observe une différence particulièrement importante du nombre de décès entre les groupes Tamoxifen et placebo au cours des 10 premières années de suivi (réponse B). En ce qui concerne la cause de décès, la proportion de décès par cancer du sein était plus importante dans le groupe Tamoxifen (réponse D), surtout pour les durées de suivi supérieure à 10 ans (réponse E). Les décès par cancer du poumon étaient moins fréquents dans le groupe Tamoxifen (réponse C). Cette étude est un bon exemple de la limite de l'utilisation des critères de jugement de substitution dans un essai de prévention. L'enregistrement de données de suivi à long-terme reste fondamentale dans ce type d'essai pour confirmer que l'impact sur le critère de substitution entraîne bien, à la fin, un impact positif sur l'événement que l'on cherche à prévenir, à savoir ici la mortalité par cancer du sein.

**Question 14**

**Réponses exactes : A – B – D**

*Commentaires* L'absence de démonstration d'un effet protecteur du Tamoxifen sur la mortalité n'est pas due à un manque de puissance statistique. L'augmentation de la taille de l'échantillon, n'aurait probablement conduit qu'à une amélioration de la précision des estimations, et donc à la démonstration d'un effet délétère statistiquement significatif du Tamoxifen sur le risque de décès.

**Question 15**

**Réponse exacte : A**

*Commentaires* Cette patiente n'est pas à « haut risque » de cancer de cancer du sein tel que défini dans cette étude. Les résultats de cet essai de prévention ne s'appliquent donc pas à elle (réponse A).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 15 Essai de prévention

Nelly Agrinier, Aurélie Bannay

| | |
|---|---|
| **Titre de l'article** | **Metabolic Syndrome From Adolescence to Early Adulthood. Effect of Infancy-Onset Dietary Counseling of Low Saturated Fat: The Special Turku Coronary Risk Factor Intervention Project (STRIP)** |
| **Auteurs** | Nupponen M ; Pahkala K ; Juonala, M ; Magnussen CG ; Niinikoski H ; Zönnemaa T ; Viikari VSA ; Saarinen V ; Lagström H ; Jula A ; Simell O ; Raitakari OT |
| **Référence du n° de la revue** | Circulation. 2015 ; 131:605-613 |
| **Lien vers le site** | **http://circ.ahajournals.org/content/131/7/605.abstract** |

**1** **Quels les sont les mots-clés principaux caractérisant l'étude STRIP ?**

a. Supériorité
b. Randomisée
c. Expérimentale
d. En double insu
e. Cross-over

**2** **Les auteurs précisent que l'étude comprenait les 534 participants qui avaient des données complètes pour la mesure du syndrome métabolique. Quelles conséquences cela peut-il avoir sur les résultats ?**

a. Biais d'attrition
b. Biais de mesure sur le critère de jugement
c. Biais de sélection
d. Perte de puissance
e. Augmentation du risque de deuxième espèce

**3** **Quels sont les éléments qui caractérisent les modalités de mesure du poids dans cette étude ?**

a. Précises
b. Claires
c. Reproductibles
d. Sujettes au risque d'erreur aléatoire
e. Sources de biais majeur



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **Concernant les choix retenus sur le critère de jugement principal de cette investigation, diriez-vous que**

a. Il présente une définition consensuelle unique chez l'adolescent
b. En tout, 5 définitions différentes ont été utilisées
c. La sensibilité des résultats à la définition utilisée a été rapportée
d. Il y a un risque augmenté de conclure à tort à l'efficacité de l'intervention
e. C'est un critère intermédiaire

**5** **Concernant les perdus de vue parmi les 534 individus inclus dans cette investigation, diriez-vous que**

a. Il n'y a pas de perdus de vue
b. Il y a autant de perdus de vue dans le groupe intervention que dans le groupe de référence
c. Il y a significativement plus de perdus de vue dans le groupe intervention que dans le groupe de référence
d. Il y a davantage de perdus du vue à l'âge de 18 ans qu'à l'âge de 20 ans
e. Le statut perdu de vue est lié à l'intervention

**6** **Concernant la comparabilité des adolescents des groupes STRIP et de référence inclus dans les analyses, diriez-vous que**

a. La distribution du sexe est comparable entre les 2 groupes
b. Il y a significativement plus de filles dans le groupe de référence
c. Les caractéristiques socio-économiques présentées permettent d'affirmer que les groupes sont comparables sur ces critères à l'inclusion
d. La randomisation à l'âge de 6 mois des adolescents inclus assure la comparabilité des 2 groupes
e. L'inclusion de la variable sexe dans les modèles statistiques étudiant l'association entre intervention et syndrome métabolique est justifiée pour limiter un biais de confusion

**7** **Concernant la prévalence du syndrome métabolique présentée dans la figure de l'article, diriez-vous que**

a. Quel que soit l'âge, la fréquence est supérieure chez les individus du groupe de référence
b. La fréquence ne semble pas varier avec l'âge dans le groupe recevant l'intervention STRIP
c. La fréquence semble varier avec l'âge dans le groupe de référence
d. La figure suggère que l'intervention est plus efficace sur la fréquence chez les patients évalués à 18 ans que chez les patients évalués à 15 ans
e. La figure suggère un effet bénéfique de l'intervention quel que soit l'âge



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**8** **Concernant l'estimation de l'effet de l'intervention STRIP sur le risque de syndrome métabolique présenté dans le texte, diriez-vous que**

a. L'intervention est bénéfique sur le risque de syndrome métabolique
b. L'intervention diminue le risque de syndrome métabolique de 59 % selon le premier RR présenté
c. La significativité du résultat est insensible au changement de définition du syndrome métabolique
d. L'interaction du groupe avec l'âge est significative
e. L'interaction du groupe avec le sexe est significative

**9** **Concernant l'effet de l'intervention STRIP sur les triglycérides (l'une des composantes du syndrome métabolique ; texte et Tableau 2, colonne ≥80th), diriez-vous que**

a. L'effet de l'intervention n'était pas le même chez les garçons et les filles
b. Le risque d'hypertriglycéridémie chez les garçons était égal au risque d'hypertriglycéridémie chez les filles multiplié par 0,71
c. Le risque d'hypertriglycéridémie chez les garçons exposés à STRIP était diminué de 29 % par rapport au groupe de référence
d. L'effet de l'intervention STRIP était significativement délétère chez les filles en absence de confusion résiduelle
e. L'effet de l'intervention STRIP était significativement bénéfique chez les garçons en absence de confusion résiduelle

**10** **À partir des résultats et des tableaux 2 et 3, on peut dire que cette étude a montré un effet significatif de l'intervention STRIP sur**

a. L'obésité abdominale (périmètre abdominal élevé)
b. L'hyperglycémie quel que soit l'âge
c. L'hypertension artérielle définie par une tension artérielle supérieure ou égale au 80e percentile
d. L'hypertension artérielle définie par une tension artérielle supérieure ou égale au 90e percentile
e. Le défaut de cholestérol HDL quel que soit le seuil de cholestérol retenu dans la définition

**11** **Parmi les composantes du syndrome métabolique, quelle paire de composantes était la plus corrélée ?**

a. La tension artérielle systolique et le tour de taille dans le groupe intervention STRIP
b. La tension artérielle systolique et la tension artérielle diastolique dans le groupe intervention STRIP
c. Le cholestérol HDL et le tour de taille dans le groupe intervention de référence
d. La tension artérielle diastolique et les triglycérides dans le groupe intervention STRIP
e. Le glucose et le tour de taille dans le groupe intervention STRIP



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**12** **Concernant la corrélation du tour de taille avec le glucose, diriez-vous que**

a. Le coefficient de corrélation est statistiquement plus élevé dans le groupe de référence que dans le groupe intervention STRIP
b. Les coefficients de corrélation sont équivalents dans le groupe de référence et dans le groupe intervention STRIP
c. Le résultat observé peut être lié à une inflation du risque de première espèce
d. Le résultat observé peut être lié à un manque de puissance
e. Le résultat observé peut être lié à un biais de confusion

**13** **L'intervention évaluée comprenait les éléments suivants**

a. Des conseils diététiques visant à diminuer les apports de sel dans l'alimentation
b. Des conseils diététiques visant à favoriser les apports de sucres lents dans l'alimentation
c. Des conseils diététiques visant à favoriser les apports de fruits et légumes
d. Des menus équilibrés qui étaient livrés régulièrement livrés à domicile pour les enfants
e. Des dépliants indiquant comment améliorer la pratique d'une activité physique régulière

**14** **Concernant l'intervention réellement reçue par les participants, diriez-vous que**

a. Les résultats présentés permettent de s'assurer de l'observance des individus du groupe STRIP
b. Les méthodes choisies permettent de s'assurer de l'absence de contamination dans le groupe de référence
c. L'intervention est suffisamment bien décrite pour être reproduite dans une autre population
d. Les résultats présentés permettent de s'assurer que les participants du groupe de référence ont été suivis en pratique comme le prévoyait le protocole
e. Les résultats présentés ne permettent pas de s'assurer que l'effet observé sur le syndrome métabolique est dû à l'intervention STRIP

**15** **Quels sont le ou les arguments apportés par l'étude qui plaident en faveur de la mise en œuvre de l'intervention évaluée ?**

a. Les résultats principaux en faveur d'un effet bénéfique de l'intervention sur le syndrome métabolique
b. L'absence de sensibilité des résultats au changement de définition du syndrome métabolique
c. La plausibilité de l'effet de l'intervention sur les critères de jugements utilisés
d. La réalisation d'analyses en intention de traiter
e. La force de l'association entre l'intervention STRIP et le critère de jugement principal



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 1

**Réponses exactes : A – B – C**

**Indispensables : B – C**

*Commentaires* Toutes les études randomisées sont des études dans lesquelles l'investigateur affectent aléatoirement les personnes incluses aux différents groupes d'exposition à l'intervention, en l'occurrence le groupe recevant l'intervention STRIP ou celui recevant l'intervention de référence. Par extension, quand l'exposition à l'intervention évaluée est maîtrisée (décidée) par l'investigateur, on qualifie l'étude d'expérimentale, par opposition aux études observationnelles dans lesquelles l'investigateur peut seulement mesurer le niveau d'exposition à l'intervention ou au facteur d'exposition étudié.

Les éléments justifiant ces réponses se trouvent :

- en fin d'introduction : « the intervention may reduce the clustering of overweight-related cardiometabolic risk factors between the ages of 5 to 15 years, suggesting a beneficial effect on the risk of MetS », et « we have now tested the hypothesis that the infant-onset intervention *reduces the risk* of pediatric MetS », en faveur de l'hypothèse de supériorité de l'intervention STRIP évaluée par rapport à l'intervention de référence. Même si le risque attendu est inférieur chez ceux qui sont exposés à l'intervention, on qualifie l'étude de « supériorité » car le résultat attendu en termes de santé produite est meilleur chez ceux qui bénéficient de l'intervention ;
- au début de la partie « Méthodes » : « 1062 infants (56,5 % of the eligible age cohort) were *randomly* allocated to an intervention (n = 540) or a control (n = 522) group. »

L'étude n'était pas en double insu. En effet dans un essai de prévention où un régime alimentaire est évalué, il est impossible de tenir le patient en aveugle de l'intervention.

Enfin, il n'y avait pas de cross-over. Les participants ne recevaient pas l'une puis l'autre intervention (STRIP puis référence OU référence puis STRIP). Cela aurait été impossible ici étant donné la durée de suivi nécessaire pour obtenir le critère de jugement (15 à 20 ans). Par ailleurs, il eût été impossible d'assurer une période de wash out, les parents des enfants du groupe « STRIP puis référence » auraient probablement pris des bonnes habitudes alimentaires pour leur enfant en lien avec les conseils diététiques reçus en période STRIP.

## Question 2

**Réponses exactes : A – C – D – E**

*Commentaires* En début de paragraphe « Méthodes », les auteurs précisent que 1062 enfants ont été randomisés (« 1062 infants (56,5 % of the eligible age cohort) were randomly allocated to an intervention (n = 540) or a control (n = 522) group ») et donc inclus dans l'essai STRIP. Seulement 534 (50 %) ont finalement été inclus dans les analyses. De plus ces 534 enfants n'ont pas été sélectionnés aléatoirement, mais sur le fait de présenter des données manquantes sur les variables permettant de définir le critère de jugement principal. Il y a donc parmi eux des perdus de vue (totalité des données manquantes au temps de suivi), et des enfants avec une partie des données manquantes au temps de suivi. Dans un essai de supériorité, l'analyse doit être réalisée en intention de traiter, c'est-à-dire que l'ensemble des participants randomisés doit être inclus dans les analyses pour assurer la comparabilité des groupes (STRIP et référence en l'occurrence). A défaut, il y a un risque de biais d'attrition, forme particulière de biais de sélection. Il est spécifié dans cet article que l'analyse a été réalisée en intention de traiter (dans le paragraphe « Analyses statistiques », dans l'extrait « Participants with missing visits in the sequence were included in the analyses (intention-to-treat analysis) », ce qui est inexact puisque, comme indiqué ci-dessus, une analyse en intention de traiter repose sur l'inclusion



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

de l'ensemble des participants randomisés (ici n = 1062) et non pas seulement une partie de ces participants sélectionnés sur le fait d'avoir l'information sur l'intégralité des composants du critère de jugement pour au moins l'un des temps de suivi (n = 534).

Enfin, toute diminution de la taille de l'échantillon dans une étude comparative aboutit à une perte de puissance et à un défaut de précision des estimations de mesure d'association entre l'exposition et le critère de jugement.

### Question 3

**Réponses exactes : A – B – C – D**

*Commentaires* Les modalités de mesure du poids sont décrites dans le paragraphe « Méthodes », où il est précisé que « Weight was measured to the nearest 0,1 kg with an electronic scale (S10, Soehnle, Murrhardt, Germany) ». Elles comprennent le type et la marque du pèse-personne, et la précision de la mesure. On aurait pu préciser les conditions de la mesure, comme pour la tension artérielle, c'est-à-dire quels vêtements étaient tolérés, mais globalement on peut qualifier les modalités de mesure de précises, claires, et reproductibles. Si l'on devait refaire la même étude on saurait comment s'y prendre pour mesurer le poids.

Par ailleurs toute mesure est sujette à un risque d'erreur aléatoire. En revanche on ne peut pas affirmer que la mesure du poids sera sujette à une erreur systématique majeure sur les éléments dont on dispose. On peut à la limite suspecter un biais si le pèse-personne n'était pas calibré au moment de son installation ou après tout déplacement.

### Question 4

**Réponses exactes : C – D – E**

*Commentaires* Les définitions pour le critère de jugement sont décrites dans le paragraphe intitulé « Defintion of MetS » dans la partie « Méthodes ». Il n'y a pas de consensus sur la définition du syndrome métabolique chez l'adolescent comme indiqué dans le début du paragraphe « there is no single or universally accepted definition of adolescent MetS ». Celles retenues par les auteurs sont au nombre de 5 binaires (celles qui permettent de classer les individus en ayant le syndrome ou ne l'ayant pas). La 6e définition présentée est celle d'un score de risque continu de syndrome métabolique. Il y a donc au total 6 définitions utilisées au total. Le fait de retenir des définitions différentes en cas d'absence de définition consensuelle est un gage de qualité, cela permet de vérifier que les résultats restent les mêmes quelle que soit la définition appliquée (reproductibilité), ou en d'autres termes de tester la sensibilité des résultats aux changements de définition. En revanche, les définitions retenues n'étant pas indépendantes entre elles (elles reposent sensiblement sur les mêmes variables) et les analyses étant réalisées sur le même échantillon, les résultats des tests statistiques réalisés devront être interprétés au regard de cette dépendance. Ainsi en modifiant les définitions sans ajuster le risque de première espèce (c'est-à-dire le seuil qui permet de conclure à la significativité des résultats), on augmente la probabilité de conclure à l'efficacité de l'intervention, initialement fixée à 5 %. C'est pourquoi on parle d'inflation du risque alpha, c'est-à-dire d'une augmentation de la probabilité de rejeter l'hypothèse nulle à tort (ou encore de conclure à un effet bénéfique alors qu'en réalité cet effet n'existe pas). Il y a des techniques utilisées pour tenir compte de cette inflation dans les analyses statistiques. La technique de Bonferroni décrite dans le paragraphe analyses statistiques en est une. Cependant elle est utilisée ici seulement dans l'étude des interactions et non pas dans les analyses de sensibilité décrites par les auteurs sur le critère de jugement principal.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Enfin, le critère de jugement utilisé dans l'étude est le syndrome métabolique qui est composé de variables caractérisant des facteurs de risque cardiovasculaire et non pas des évènements cardiovasculaires à proprement parler (décès de cause cardiovasculaire, syndrome coronarien aigu, accident vasculaire cérébral, etc.). Il s'agit donc bien d'un critère intermédiaire. On cherche à évaluer l'effet d'une intervention sur des facteurs de risque et non pas sur le risque de survenue d'un évènement.

### Question 5

**Réponses exactes : C – E**

**Indispensable : C**

*Commentaires* La phrase permettant de répondre à cette question se trouve dans le paragraphe « Lost to follow-up » dans la partie « Méthodes » : « The proportion of premature discontinuation was higher in the intervention group than in the control group (29 % *versus* 19 % ; P = 0,008, Cox regression model) ». 29 % > 19 % donc il y a plus de perdus de vue dans le groupe intervention que dans le groupe de référence ; p < 0,05, risque de première espèce tel que défini à la fin des analyses statistiques, donc le résultat est statistiquement significatif.

### Question 6

**Réponse exacte : A**

**Indispensable : A**

*Commentaires* Dans la partie Méthodes il est indiqué « The sex distribution was similar within the study groups (at 15 years of age: intervention group, 47 % girls; control group, 49 % girls ; P = 0,57, x 2) », la distribution du sexe est donc comparable entre les 2 groupes. Les proportions de filles à l'âge de 15 ans ne diffèrent pas entre les 2 groupes (p > 0,05). En d'autres termes, la variable sexe n'est pas associée au groupe d'intervention, et n'est donc pas un facteur de confusion dans la relation entre l'intervention et le risque de syndrome métabolique puisqu'un facteur de confusion est par définition associé à la fois au facteur d'exposition, ET au critère de jugement. Il n'y a donc pas de justification d'ajuster sur le sexe dans les modèles étudiant l'association entre l'intervention et le syndrome métabolique.

### Question 7

**Réponses exactes : A – B – C – D – E**

*Commentaires* Dans la figure, quel que soit l'âge les barres blanches, représentant la fréquence du syndrome métabolique dans le groupe de référence, sont plus hautes que les barres noires, représentant la fréquence de survenue du syndrome métabolique dans le groupe intervention. Par ailleurs les barres noires sont à peu près toutes de même hauteur (la fréquence du syndrome métabolique dans le groupe intervention varie entre 6 % et 8 %, selon l'âge, ce qui est une faible variation). En revanche les barres blanches ont une hauteur qui augmente puis qui diminue quand l'âge avance (la fréquence du syndrome métabolique varie entre 10 et 14 ou 15 %, ce qui représente une variation relative importante selon l'âge, mais non monotone (elle augmente puis elle diminue, ni linéaire, l'augmentation ne dessine pas une droite sur le graphique). En conséquence le rapport entre les fréquences de syndrome métabolique dans le groupe intervention et dans le groupe de référence devrait diminuer (se rapprocher de zéro et s'éloigner de 1) puis augmenter (s'éloigner de zéro et se rapprocher de 1) avec l'âge qui avance, en d'autres termes le rapport de risque devrait varier selon l'âge, ou l'effet de l'intervention devrait varier selon l'âge. C'est ce qui définit une interaction. Elle peut être synergique (l'effet augmente – le rapport de risque s'éloigne de 1 – avec l'âge) ou antagoniste (l'effet diminue – le



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

rapport de risque se rapproche de 1, voire s'inverse- avec l'âge), ici ce serait un mélange des 2 et ça pourrait expliquer que le terme d'interaction groupe d'intervention par âge soit non significatif (l'effet synergique intervention-âge) observé pour les âges allant de 15 à 17 ans puis l'effet antagoniste au-delà de 17 ans s'annulent quand on considère l'interaction globalement, ce qui est fait dans les modèles).

Concernant l'hypertension artérielle, les résultats en rapport avec les items C et D sont présentés dans le tableau 2 à la ligne « High blood pressure », colonne « ≥ 80th » pour l'item C et « ≥ 90th » pour l'item D. Dans le cas de l'item C on a bien un RR < 1 et un intervalle de confiance ne comprenant pas la valeur 1, donc statistiquement significatif. Dans le cas de l'item D, on a un RR < 1 mais un intervalle de confiance qui comprend la valeur 1, donc statistiquement non significatif.

## Question 8

**Réponses exactes : A – C**

**Indispensable : A**

*Commentaires* Cette question reprend les éléments du paragraphe « Effect of intervention on MetS » dans la partie « Résultats ». Quelle que soit la définition utilisée pour le syndrome métabolique, le risque relatif est inférieur à 1 (effet bénéfique) et l'intervalle de confiance ne comprend pas la valeur 1 (le résultat est statistiquement significatif). Donc l'intervention est bénéfique sur le risque de survenue de syndrome métabolique. Le premier RR présenté est à 0,59, on a donc, une incidence de syndrome métabolique dans le groupe intervention (IMetS1) qui est 0,59 fois plus faible que dans le groupe de référence (IMetS0), ou, en d'autres termes comme indiqué dans le texte, une incidence diminuée de 41 %.

RR = IMetS1/IMetS0 = 0,59 ; donc IMestS1 = 0,59 x IMetS0 ; donc IMetS1 = IMetS0 x (1-0,41)

Le terme d'interaction du groupe avec l'âge (ou avec le sexe) présente un degré de signification p = 0,95 (ou p = 0,80) (« There was no study group–by–sex (P = 0,80) or study group–by–age (P = 0,95) interactions, indicating that the intervention effect was similar among girls and boys and at different ages » ; donc il est non significatif. Néanmoins, comme suspecté dans la figure, l'effet semble varier avec l'âge de manière synergique avant 17 ans puis antagoniste après 17 ans. Ainsi l'effet de l'interaction n'est pas linéaire et le modèle linéaire utilisé ne peut le mettre en évidence (globalement on dira que l'effet ne varie pas avec le temps alors qu'en réalité il varie probablement dans 2 directions opposées sur 2 sous-groupes différents). Il est donc probable que degré de signification du terme d'interaction groupe par âge ne soit pas interprétable.

## Question 9

**Réponses exactes : A – C – E**

*Commentaires* Pour le critère de jugement hypertriglycéridémie, les auteurs expliquent que l'interaction groupe d'intervention par sexe était statistiquement significative (p < 0,05) : « There was a significant study group–by–sex interaction in triglycerides (P = 0,006) », en conséquence l'effet de l'intervention STRIP varie selon le sexe, ce qui se confirme sur les valeurs ponctuelles des RR rapportés dans le tableau 2 qui n'ont pas les mêmes valeurs selon le sexe : chez les garçons (RR < 1) et chez les filles (RR > 1) (le changement de sens de l'association n'est pas obligatoire en cas d'interaction, en revanche le changement de valeur du RR l'est). Le problème posé par ces résultats selon le sexe est liée à l'analyse en sous-groupes dans un essai : on n'est plus en mesure d'assurer la comparabilité des groupes ayant reçu ou pas l'intervention (l'allocation de l'intervention n'a pas été faite au hasard dans le sous-groupe des garçons puis dans le sous-groupe des filles, elle a été réalisée aléatoirement globalement, sans tenir compte du sexe). Il n'est donc pas exclu qu'il existe un biais de confusion résiduel (c'est pourquoi les 2 derniers



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

items précisent « en l'absence de confusion résiduelle »). Si l'on peut dire qu'en l'absence de confusion résiduelle, l'effet de l'intervention STRIP est bénéfique chez les garçons (RR < 1 et intervalle de confiance du RR ne comprend pas la valeur 1), on ne peut pas affirmer qu'il est délétère chez les filles (car même si le RR est bien supérieur à 1, l'intervalle de confiance du RR comprend la valeur 1, le résultat n'est donc statistiquement pas significatif).

Par ailleurs, concernant l'item B, il n'y a pas d'éléments quantitatifs dans les résultats qui indiquent le rapport de risque d'hypertriglycéridémie selon le sexe (ou l'estimation de l'effet du sexe sur le risque de survenue d'hypertriglycéridémie).

Enfin, comme indiqué dans le commentaire relatif à la question 8, quand le RR = 0,71, alors 1-0,71 = 0,29, et on peut dire que le risque est diminué de 29 %. Ce qui explique pourquoi l'item C est juste.

## Question 10

**Réponse exacte : C**

**Indispensable : C**

*Commentaires* Les auteurs proposent une évaluation de l'effet de l'intervention sur chacune des composantes du syndrome métabolique, c'est-à-dire d'utiliser des critères de jugement secondaires, en utilisant plusieurs définitions pour chaque critère de jugement secondaire (analyse de sensibilité), ce qui explique les 3 paires de colonnes (RR et intervalle de confiance à 95 %) dans le tableau 2. Concernant l'obésité abdominale (ligne « High waist »), tous les RR sont < 1, ce qui plaide en faveur d'un effet bénéfique de l'intervention, en revanche tous les intervalles de confiance comprennent la valeur 1, ce qui implique que l'effet n'est pas statistiquement significatif. Idem pour l'hyperglycémie (ligne « High glucose »). Il y a une subtilité pour l'hyperglycémie qui se trouve dans le texte de a partie « Résultats » : « A study group–by–age interaction on glucose indicated thatthe effectof intervention was different by age (P = 0,016). In pairwise comparisons, the risk for high glucose was lower in intervention participants at 18 years of age (RR, 0,60 ; 95 % CI, 0,38–0,96) ». L'interaction mise en évidence indique que l'effet de l'intervention diffère selon l'âge, et que chez les adolescents de 18 ans, l'effet de l'intervention sur l'hyperglycémie semblait bénéfique (RR < 1, et l'intervalle de confiance ne comprend pas la valeur 1). Ce résultat étant à interpréter avec prudence puisque les analyses ne sont pas conduites sur l'intégralité des enfants randomisés (n = 1062), et ce serait particulièrement gênant si par exemple des enfants inclus dans l'essai étaient morts avant l'âge de 15 ans en lien avec un diabète (ces enfants auraient eu des données incomplètes sur le syndrome métabolique à l'âge de 15 ans et auraient donc été exclus des analyses présentées dans l'article alors même que l'évènement ayant causé leur exclusion était peut-être en lien avec leur alimentation et donc avec l'intervention).

## Question 11

**Réponse exacte : B**

**Indispensable : B**

*Commentaires* Les résultats permettant de répondre à cette question sont exposés dans le tableau 4. La valeur d'un coefficient de corrélation varie entre -1 et 1. La corrélation est d'autant plus forte entre 2 paramètres quantitatifs continus que la valeur du coefficient de corrélation s'éloigne de zéro ou se rapproche de valeur absolue de 1 (c'est-à-dire de -1 ou de 1).

Le coefficient de corrélation le plus fort dans le tableau 4 est donc celui entre la tension artérielle systolique et la tension artérielle diastolique dans le groupe intervention (=0,55). Cliniquement, il n'est pas étonnant que les 2 paramètres soient corrélés.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 12**

**Réponses exactes : A – C – E**

**Indispensable : A**

**Inadmissible : B**

*Commentaires* Les résultats permettant de répondre à cette question sont exposés dans le tableau 4, ligne « Glucose », colonne « Waist ». Il est indiqué que le coefficient de corrélation glucose-Tour de taille est égal à 0,055 dans le groupe intervention STRIP et à 0,24 dans le groupe de référence, avec un degré de signification pour la comparaison des 2 à 0,025. Ce dernier étant inférieur à 0,05 et 0,24 > 0,055, donc l'item A est juste, et l'item B est faux.

Le fait de réaliser plusieurs comparaisons de coefficients de corrélation à partir d'un même jeu de données augmente les probabilités de conclure à une différence sur l'une des comparaisons alors qu'en réalité, cette différence n'existe pas et qu'elle est liée au hasard (inflation du risque de première espèce).

Le résultat étant statistiquement significatif, il n'y a pas de défaut de puissance. Un défaut de puissance étant, dans une hypothèse de supériorité, défini par une impossibilité de conclure à la supériorité de l'un des coefficients de corrélation par rapport à l'autre en lien avec une taille d'échantillon insuffisante.

Enfin, les comparaisons des coefficients de corrélation étant réalisées 2 à 2, elles ne tiennent pas comptent d'éventuels facteurs de confusion. Il est donc possible de mettre en évidence une corrélation entre le glucose et le tour de taille qui est due à un phénomène de confusion. Il est également possible qu'un facteur de confusion masque l'une des différences de corrélations étudiées (c'est-à-dire que si l'on tenait compte de ce facteur on parviendrait à trouver une différence entre les coefficients de corrélations que l'on ne parvient pas à mettre en évidence si l'on n'en tient pas compte).

**Question 13**

**Réponses exactes : A – B – C**

*Commentaires* Les éléments permettant de répondre à cette question se trouvent dans le paragraphe « Study Design and Participants » dans « Méthodes ». L'intervention comprenait un ensemble de conseils hygiéno-diététiques délivrés 2 fois par an jusqu'à l'âge de 20 ans comprenant : « how to reduce salt intake and to favor whole-grain products, fruit, and vegetables ». Concernant l'activité physique, l'intervention n'était pas standardisée (on ne parle nulle part de dépliant) : « A physically active lifestyle was encouraged, but itwas not a structured, continuous part of the intervention. ». A aucun moment il est question d'intervenir sur les repas réellement délivrés aux enfants, il s'agissait de conseils personnalisés : « A fixed diet was never ordered ; the counseling was individualized, and the child's recent food record was used as a basis of suggestions for dietary changes ».

**Question 14**

**Réponse exacte : E**

**Indispensable : E**

*Commentaires* Il n'y a dans l'article aucun élément présenté dans les résultats concernant l'évaluation de processus de la mise en place de l'intervention. C'est-à-dire par rapport à ce qui était prévu, qu'est-ce que les enfants ont réellement reçu ? Y a-t-il eu contamination ? C'est fort probable dans un essai de prévention avec randomisation individuelle (c'est-à-dire qu'en fonction de l'enfant inclus qui était reçu en consultation, les praticiens du centre devraient prodiguer ou pas les conseils hygiéno-diététiques personnalisés). Néanmoins, la contamination



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

devrait diluer l'effet de l'intervention, donc ce n'est pas l'élément le plus préoccupant méthodologiquement ici. Ne sachant pas ce que les enfants du groupe intervention ont réellement reçu comme conseil, on ne sait pas si c'est le fait d'être vu au moins 2 fois par an jusqu'à 20 ans qui compte ou bien la nature des conseils délivrés. En l'absence d'éléments décrivant ce que les enfants du groupe de référence ont réellement reçu on peut imaginer que certains n'ont pas été vus entre la naissance et leurs 15-20 ans, même si ce cas extrême est peu probable. Par ailleurs les analyses n'ont pas contrairement à ce qui est indiqué étaient réalisées en l'intention de traiter (voir commentaire question 2), les résultats peuvent donc souffrir d'un biais de confusion. Les recommandations CONSORT préconisent de présenter des éléments en faveur de l'appréciation de la qualité de la randomisation et de la comparabilité des groupes (comparaison des caractéristiques à l'inclusion des enfants inclus dans les analyses selon le groupe d'intervention) qui ne sont pas présentées ici, ce qui fait douter encore plus de l'absence de biais de confusion. La seule manière de s'assurer que des résultats concluant dans un essai sont attribuables à l'intervention évaluée est l'analyse en intention de traiter sur le critère de jugement principal qui conduit à un résultat significatif en faveur de l'intervention.

## Question 15

**Réponses exactes : A – B – C – E**

*Commentaires* L'intervention évaluée est bénéfique sur le critère de jugement principal (RR < 1 et intervalle de confiance ne comprenant pas la valeur 1) quelle que soit la définition retenue (voir explications question 8). Les résultats sont donc insensibles à la définition du syndrome métabolique. La plausibilité de l'effet de l'intervention est discutée dans la partie « Discussion » : « Reasons for the intervention effect may relate to the tendency for a lower risk for high waist circumference in the intervention group and dietary factors, for example, quality of fat, 8,25,26 fiber intake, 10,27 and vegetable/ potassium intake, 9,28,29 which are shown to be associated with blood pressure and are affected by the intervention ». La taille de l'effet observée est importante, un RR = 0,71 signifie que le risque de syndrome métabolique chez les enfants ne recevant pas l'intervention STRIP est multiplié par (1/0,71) = 1,4. Enfin comme indiqué en commentaire de la question 2 les analyses dont les résultats ont été présentés dans cet article ne sont pas réalisées en intention de traiter. L'étude réalisée s'apparente plus à une analyse en sous-groupe (sous-groupe des enfants inclus dans l'essai ayant des données complètes sur le critère de jugement principal à l'un des temps de suivi) ou à une cohorte nichée dans un essai.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Études observationnelles

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 1 Étude de cohorte

Lukshe Kanagaratnam

| | |
|---|---|
| **Titre de l'article** | **Clinical deterioration in older adults with delirium during early hospitalisation: a prospective cohort study** |
| **Auteurs** | Hsieh SJ, Madahar P, Hope AA, Zapata J, Gong MN |
| **Référence** | BMJ Open. 2015 ; 5(9) : e007496 |
| **Lien vers l'article** | **https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4567670/** |

**1** **Les objectifs de cette étude étaient**

a. De déterminer la prévalence des syndromes confusionnels aigus chez les sujets âgés durant les trois premiers jours d'hospitalisation
b. De déterminer l'incidence des syndromes confusionnels aigus chez les sujets âgés durant les trois premiers jours d'hospitalisation
c. De décrire la durée des syndromes confusionnels aigus diagnostiqués chez les sujets âgés durant les trois premiers jours d'hospitalisation
d. D'identifier les facteurs associés à une ré-hospitalisation précoce
e. D'identifier les facteurs associés au déclin cognitif

**2** **Il s'agit d'une étude**

a. De cohorte
b. Transversale
c. Longitudinale
d. Analytique
e. Prospective

**3** **Les patients inclus étaient**

a. Des patients âgés de plus de 65 ans
b. Les patients admis aux urgences entre juillet 2011 et novembre 2011
c. Les patients consultant aux urgences entre 7 heures et 19 heures
d. Les patients consultant aux urgences du lundi au samedi
e. Les patients ayant donnés leur consentement écrit

**4** **Les patients n'étaient pas inclus**

a. S'ils étaient transférés directement dans une unité de soins intensifs
b. S'ils ne parlaient pas anglais
c. S'ils n'étaient pas hospitalisés après leur passage aux urgences
d. S'ils étaient atteints d'une démence sévère
e. S'ils étaient dans le coma



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**5** **Le critère de jugement principal dans cette étude était**

a. Un critère composite
b. La détérioration clinique pendant l'hospitalisation
c. L'insuffisance organique
d. La durée d'hospitalisation
e. L'institutionnalisation

**6** **Le syndrome confusionnel aigu**

a. A été évalué par des étudiants en médecine
b. A été évalué par des internes en médecine
c. A été évalué à partir de l'outil CAM-ICU
d. A été évalué à partir de l'échelle RASS
e. A été recherché pendant toute la durée de l'hospitalisation

**7** **D'après la figure 1**

a. 435 patents ont été inclus
b. 260 patients ont été analysés
c. 14 patients ont été exclus car ils étaient dans le coma
d. 26 patients ont été exclus car ils étaient hospitalisés dans une unité de soins intensifs
e. 135 patients ont refusé de participer à l'étude

**8** **Concernant la population étudiée**

a. La durée moyenne de séjour était de 5 jours
b. 38 patients ont présentés un épisode de syndrome confusionnel aigu durant les trois premiers jours d'hospitalisation
c. 156 sujets inclus étaient des femmes
d. L'âge moyen des patients inclus étaient de 85 ans
e. Le score REMS moyen était de 8

**9** **Les patients ayant présentés au moins un épisode de syndrome confusionnel aigu durant les trois premiers jours de leur hospitalisation**

a. Étaient significativement plus âgés que les patients sans syndrome confusionnel
b. Avaient un score REMS significativement plus élevé que les patients sans syndrome confusionnel
c. Avaient plus souvent des problèmes respiratoires
d. Étaient plus souvent admis à l'hôpital pour une infection
e. Étaient plus souvent admis pour un problème neurologique



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**10** **D'après les résultats présentés dans le tableau 2**

a. Il y avait peu de patients avec une détérioration clinique
b. Il y avait significativement plus de décès durant l'hospitalisation dans le groupe de patients ayant présentés un syndrome confusionnel comparé au groupe de patients sans syndrome confusionnel
c. Il y avait significativement plus de passage en unités de soins intensifs dans le groupe de patients ayant présentés un syndrome confusionnel comparé au groupe de patients sans syndrome confusionnel
d. Il y avait significativement plus d'insuffisance cardiaque dans le groupe de patients ayant présentés un syndrome confusionnel comparé au groupe de patients sans syndrome confusionnel
e. Il y avait significativement plus de patients institutionnalisés dans le groupe de patients ayant présentés un syndrome confusionnel comparé au groupe de patients sans syndrome confusionnel

**11** **La figure 2 présente**

a. Des analyses stratifiées sur la présence d'un syndrome confusionnel
b. Des analyses stratifiées sur le statut vital
c. Des analyses stratifiées sur l'âge et le score REMS
d. Des analyses multivariées par une régression logistique
e. Des analyses multivariées par un modèle de Cox

**12** **Concernant les analyses sur le critère de jugement principal**

a. Une analyse multivariée a été effectuée après ajustement sur l'âge et le score REMS
b. En analyse multivariée, le risque de détérioration clinique était significativement plus élevé dans le groupe de patients avec un syndrome confusionnel aigu comparé au groupe de patients sans syndrome confusionnel
c. L'odds ratio ajusté était de 10,21
d. L'intervalle de confiance à 95 % de l'odds ratio ajusté était de [1,91-34,14]
e. L'intervalle de confiance à 95 % de l'odds ratio ajusté était large car il y avait peu d'événements

**13** **Les limites de cette étude étaient**

a. La présence d'un biais de classement
b. La présence d'un biais de sélection
c. L'utilisation de l'outil CAM-ICU pour déterminer la présence d'un syndrome confusionnel
d. Le nombre d'événement
e. L'absence de prise en compte des syndromes confusionnels survenant après les trois jours d'hospitalisation



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**14** **Les auteurs auraient dû**

a. Prolonger la période d'inclusion pour augmenter le nombre d'événements
b. Réaliser une étude multicentrique pour augmenter le nombre d'événements
c. Prendre en compte les différents facteurs de confusion comme la présence d'une dépression avant l'hospitalisation
d. Inclure les patients se présentant aux urgences entre 8 heures et minuit
e. Former les assistants de recherche à l'outil CAM-ICU

**15** **Pour conclure**

a. La présence d'un syndrome confusionnel durant les trois premiers jours d'hospitalisation est associée à un risque de détérioration clinique plus importante
b. Le risque de détérioration clinique augmente avec la durée du syndrome confusionnel
c. Le risque de détérioration clinique augmente avec l'âge du patient
d. Le syndrome confusionnel est sous-diagnostiqué chez le sujet âgé aux urgences
e. D'autres études sont nécessaires pour valider les résultats trouvés dans cette étude



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 1**

**Réponses exactes : A – B – C**

*Commentaires* Les réponses A, B, C apparaissent dans le texte à la fin de l'introduction, en revanche les réponses D, E ne sont pas citées dans le texte.

**Question 2**

**Réponses exactes : A – C – D – E**

*Commentaires* Les réponses A et E apparaissent dans la partie méthodes. La réponse C est juste car une étude de cohorte est une étude longitudinale. Il s'agit d'une étude analytique car les auteurs comparent deux groupes de patients (l'un présentant un syndrome confusionnel aigu et l'autre sans syndrome confusionnel aigu) et recherchent l'existence d'une association entre la présence d'un syndrome confusionnel aigu et la survenue d'un décès ou le passage dans une unité de soins intensifs.

**Question 3**

**Réponses exactes : A – B – C**

*Commentaires* Les réponses A, B, C apparaissent dans la partie méthodes. La réponse D est fausse car les patients inclus étaient les patients consultant aux urgences du lundi au vendredi. La réponse E est fausse car les patients inclus étaient ceux ayant donnés leur accord oral.

**Question 4**

**Réponses exactes : A – B – C – D – E**

*Commentaires* Les réponses A, B, C, D, et E apparaissent dans la partie méthodes.

**Question 5**

**Réponses exactes : A – B**

*Commentaires* Les réponses C, D, E sont fausses car elles correspondent aux critères de jugement secondaire. Le critère de jugement principal est un critère composite (il regroupe plusieurs événements cliniques) : passage dans une unité de soins intensifs ou décès à l'hôpital.

**Question 6**

**Réponses exactes : A – B – C**

*Commentaires* L'échelle RASS a été utilisée pour évaluer la conscience. Le syndrome confusionnel aigu a été recherché durant les trois premiers jours d'hospitalisation.

**Question 7**

**Réponses exactes : B – E**

*Commentaires* Toutes les réponses apparaissent dans la figure 1.

**Question 8**

**Réponses exactes : B – C**

*Commentaires* La réponse A est fausse car il s'agit de la durée médiane de séjour (la durée moyenne n'étant pas précisée dans l'article). La réponse B apparaît dans la partie résultat. La réponse C est exacte car d'après le tableau, il y a 136 femmes dans le groupe de patients n'ayant pas présentés de syndrome confusionnel et 20 dans le groupe de patients avec syndrome confusionnel. La réponse D et E sont fausses car non précisées par les auteurs dans l'article.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 9**

**Réponses exactes : A – B – D**

*Commentaires* Toutes les réponses apparaissent dans le tableau 1. Il existe une différence significative car le p est inférieur à 0,05.

**Question 10**

**Réponses exactes : A – B – C**

*Commentaires* Toutes les réponses apparaissent dans le tableau 2.

**Question 11**

**Réponse exacte : D**

*Commentaires* La figure 2 présente les résultats des analyses non ajustées et ajustées pour les critères de jugement principal et secondaire (par une régression logistique). La réponse figure dans le paragraphe « analyse des données » de la partie méthodes.

**Question 12**

**Réponses exactes : A – B – D – E**

*Commentaires* Toutes les réponses apparaissent dans la figure 2 et page 5, 2e colonne.

**Question 13**

**Réponses exactes : A – B – C – D – E**

*Commentaires* Les réponses apparaissent dans la discussion. Il existe un biais de classement lié à l'utilisation de l'outil CAM-ICU, certains patients avec un syndrome confusionnel auraient été classés dans le groupe de patients sans syndrome confusionnel. Dans cette étude, le nombre d'événement (détérioration clinique) était faible et n'a pas permis de faire une analyse multivariée en tenant des facteurs de confusion potentiels.

**Question 14**

**Réponses exactes : A – B – C**

*Commentaires* Les auteurs auraient dû prolonger la période d'inclusion ou réaliser une étude multicentrique pour augmenter le nombre de patients et donc le nombre d'événement, permettant ainsi de faire une analyse multivariée en tenant compte des facteurs de confusion. La réponse C est exacte et apparaît dans la discussion. La réponse D est fausse, les auteurs auraient dû inclure tous les patients qui se présentaient aux urgences sans tenir compte de l'heure d'admission. La réponse E est fausse car les assistants de recherche ont été formés à l'outil CAM-ICU.

**Question 15**

**Réponses exactes : A – D – E**

*Commentaires* Les réponses A, D, E apparaissent dans la partie résultat et/ou discussion. Les réponses B, C n'apparaissent pas dans l'article. La présence d'un syndrome confusionnel n'était pas renseignée dans 52 % des dossiers médicaux (page 4, 2e colonne, dernier paragraphe et page 5, 1re colonne, 1re paragraphe) Les analyses ont été ajustées sur l'âge mais les auteurs ne donnent pas la valeur de l'odds ratio ni de son intervalle de confiance à 95 % (aucune estimation n'est disponible pour ce facteur et nous ne pouvons donc pas savoir si le risque de détérioration clinique augmente avec l'âge du patient).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 2 Étude de cohorte (score de propension)

Aurélie Bannay

| | |
|---|---|
| **Titre de l'article** | **Increased risk of cardiovascular events in stroke patients who had not undergone evaluation for coronary artery disease** |
| **Auteurs** | Kim YD, Song D, Nam HS, Choi D, Kim JS, Kim BK, Chang HJ, Choi HY, Lee K, Yoo J, Lee HS, Nam CM, Heo JH |
| **Référence du n° de la revue** | Yonsei Med J. 2017 Jan ; 58(1):114-122. |
| **Lien vers le site** | **https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5122626/** |

**1** **Cette étude s'inscrit dans le champ**

a. De la prévention primaire
b. Du dépistage organisé
c. De la prévention secondaire
d. De la médecine personnalisée
e. De la prévention tertiaire

**2** **Les caractéristiques suivantes peuvent s'appliquer à cette étude**

a. Expérimentale
b. Observationnelle
c. De cohorte
d. Cas-témoins
e. Exposé-non exposés

**3** **Dans cette étude, le terme « rétrospectif » s'applique**

a. Au suivi des sujets
b. Au recrutement des sujets
c. À la survenue du critère de jugement principal
d. À la randomisation
e. À la survenue de l'exposition



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **Dans cette étude, le ou les critères de jugement**

a. Étai(n)t combinés
b. Nécessitai(en) t la survenue d'au moins un événement cardiovasculaire et d'une récidive d'AVC ischémique
c. Étai(en) t constitués d'un (ou des) code(s) diagnostique(s) issus de la 10e édition de la Classification Internationale des Maladies
d. Étai(en) t un (ou des) délai(s)
e. Étai(en) t un(ou des) score(s)

**5** **D'après vous, l'objectif de cette étude était de montrer**

a. L'effet d'une évaluation par coroscanner sur la mortalité à court terme des patients après un AVC ischémique ou un AIT sans facteur de risque cardio-vasculaire
b. L'effet d'une évaluation par coroscanner sur la récidive de syndrome coronarien aigu à long terme des patients après un AVC ischémique ou un AIT
c. L'effet d'une évaluation par coroscanner sur la survenue à long terme du décès ou d'cardiaque ou cérébro-vasculaire de patients après un AVC ischémique ou un AIT, sans antécédent de coronaropathie
d. L'effet d'une évaluation par coroscanner sur la survenue à long terme du décès ou d'cardiaque ou cérébro-vasculaire de patients après un AVC ischémique ou un AIT, avec antécédent de coronaropathie
e. L'effet d'une évaluation par coroscanner sur la survenue à court terme d'cardiaque ou cérébro-vasculaire de patients après un AVC ischémique ou un AIT, sans antécédent de coronaropathie

**6** **Concernant l'inclusion des sujets, on peut affirmer que**

a. L'exclusion des sujets ayant des antécédents de coronaropathie n'était pas pertinente
b. L'exclusion des sujets non éligibles au coroscanner était pertinente
c. L'exclusion des sujets décédés dans les 7 jours suivant l'accident vasculaire initial était pertinente
d. Les différents centres n'ont pas appliqué exactement les mêmes critères d'inclusion
e. Tous les sujets ont explicitement donné leur consentement à cette étude avant leur inclusion

**7** **Concernant l'analyse des 3 117 sujet inclus, on peut affirmer que**

a. Les sujets n'ayant pas bénéficié d'un coroscanner ont des caractéristiques à l'inclusion similaires aux sujets n'ayant pas bénéficié d'un coroscanner
b. Plus de 67,2 % des 3 117 sujets sont montré des signes de coronaropathie asymptomatique
c. La proportion de sujets diabétiques dans le groupe n'ayant pas bénéficié du coroscanner était supérieure à celle dans le groupe en ayant bénéficié
d. Durant le suivi, le nombre de décès était supérieur au nombre de récidives d'accident vasculaire cérébral et au nombre d'cardiovasculaire
e. Tous les sujets ont eu la même durée de suivi



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**8** **D'après le tableau 1 et le texte de l'article, on peut affirmer que**

a. Seuls 808 sujets ont été inclus dans l'analyse appariée sur le score de propension
b. Dans l'analyse appariée sur le score de propension, les sujets ayant bénéficié de l'évaluation par coroscanner ont des caractéristiques identiques à celles des sujets n'en ayant pas bénéficié
c. Dans l'analyse appariée sur le score de propension, les sujets ayant bénéficié de l'évaluation par coroscanner sont issus de la même population en termes de caractéristiques présentées dans le tableau 1 que les sujets n'ayant pas bénéficié de l'évaluation par coroscanner
d. Les caractéristiques des sujets appariés sont toujours situées entre celles du groupe non apparié ayant bénéficié du coroscanner et celle du groupe non apparié n'en ayant pas bénéficié
e. Deux échantillons d'analyse ont été considérés dans cet article

**9** **La figure 2 permet d'affirmer que**

a. Les sujets de l'étude qui ont bénéficié du coroscanner survivent plus longtemps que les sujets n'en ayant pas bénéficié
b. Les sujets de l'étude qui ont bénéficié du coroscanner meurent moins fréquemment d'événements cardiovasculaires que les sujets n'en ayant pas bénéficié
c. Les sujets de l'étude qui ont bénéficié du coroscanner meurent moins fréquemment de récidives d'accident vasculaire cérébral que les sujets n'en ayant pas bénéficié
d. L'évaluation par coroscanner permet d'améliorer la survie des patients après un épisode d'accident vasculaire cérébral
e. L'évaluation par coroscanner n'était pas statistiquement associée à la survenue d'événement cardiovasculaire

**10** **Dans cette étude, le score de propension**

a. Correspondait à la probabilité pour un sujet de bénéficier de l'évaluation par coroscanner à partir de ses caractéristiques initiales
b. Était égal à 1 pour les sujets ayant bénéficié de l'évaluation par coroscanner
c. Était égal à 0 pour les sujets n'ayant pas bénéficié de l'évaluation par coroscanner
d. A été obtenu par une régression logistique modélisant la probabilité de survie
e. Pouvait être égal pour un sujet ayant bénéficié de l'évaluation par coroscanner et pour un sujet n'ayant pas bénéficié de l'évaluation par coroscanner

**11** **D'après le tableau 2, concernant les modèles statistiques dans cette étude, on peut affirmer que**

a. Cinq types d'analyses basées sur le score de propension ont été menés
b. L'analyse multivariée avec un ajustement simple a donné des résultats significatifs sur les mêmes critères de jugement que l'analyse non ajustée
c. Le modèle non ajusté a montré des hazard ratio plus élevés que tous les modèles basés sur le score de propension quel que soit le critère de jugement
d. L'analyse dans chaque quintile du score de propension avait une puissance moins élevée que les autres analyses multivariées quel que soit le critère de jugement
e. Chacun des hazard ratio est issu d'un modèle de survie ajusté



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**12** **Dans le tableau 2, l'analyse dans chaque quintile du score de propension permet d'affirmer que**

a. Les sujets ayant la plus faible probabilité d'être évalués par coroscanner montraient un effet néfaste significatif de l'évaluation par coroscanner sur la survie
b. Les sujets ayant la plus forte probabilité d'être évalués par coroscanner avaient un risque de décès 2,88 fois plus élevé que les sujets ayant la plus faible probabilité d'être évalués par coroscanner
c. Chez les sujets ayant la plus forte probabilité d'être évalués par coroscanner, l'absence d'évaluation par coroscanner était associée à un risque 2 fois plus élevé de récidive d'AVC
d. Seuls les sujets ayant la plus forte probabilité d'être évalués par coroscanner montraient un bénéfice significatif de l'évaluation par coroscanner à la fois en termes de décès, d'cardiovasculaire et de récidive d'AVC.
e. Aucune association significative entre l'évaluation par coroscanner et le décès n'était retrouvée

**13** **Une analyse de sensibilité**

a. A été effectuée entre les sujets ayant bénéficié d'un coroscanner et les sujets n'en ayant pas bénéficié pour des raisons non médicales
b. A utilisé des modèles d'analyse de survie classique et basés sur le score de propension comme l'analyse principale
c. Confirmait l'association entre l'évaluation par coroscanner et le décès par une analyse basée sur le score de propension
d. Confirmait l'association entre l'évaluation par coroscanner et la survenue d'cardiovasculaire par une analyse basée sur le score de propension
e. Confirmait l'association entre l'évaluation par coroscanner et la récidive d'AVC par une analyse basée sur le score de propension

**14** **Concernant les limites de cette étude, on peut affirmer que**

a. L'absence de randomisation pouvait induire un biais de confusion
b. L'analyse basée sur le score de propension avait pour but de limiter le biais d'indication induit par l'absence de randomisation
c. Des facteurs de confusion pourraient ne pas avoir été pris en compte par l'analyse basée sur le score de propension
d. Le caractère monocentrique augmente la généralisation des résultats de l'étude
e. Le caractère monocentrique diminue la généralisation des résultats de l'étude

**15** **Concernant l'indépendance des auteurs par rapport à l'industrie, on peut affirmer que**

a. Le financement est exclusivement d'origine institutionnelle au travers d'une subvention du ministère de la santé coréen
b. La revue ayant publié cet article déclare être financée par l'industrie
c. Les médicaments prescrits durant le suivi ont été pris en charge par l'industrie
d. Le coroscanner a été mis à disposition par l'industrie
e. Les auteurs ont déclaré une absence de conflit d'intérêts



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 1

**Réponse exacte : E**

**Indispensable : E**

*Commentaires* Cette étude a inclus des sujets ayant été hospitalisés pour la prise en charge initiale d'un accident vasculaire cérébral (AVC). Elle ne s'inscrit donc pas dans le cadre de la prévention primaire qui concerne des sujets non atteints par la pathologie concernée. Elle s'intéresse à la détection de complications (événement cardiovasculaire) ou de récidive d'AVC chez des sujets malades. Elle relève donc de la prévention tertiaire. La médecine personnalisée vise à utiliser des données individuelles en particulier génétiques pour déterminer le choix du traitement ou des moyens de prévention les plus efficaces pour un patient donné ce qui n'est pas le cas dans cette étude.

### Question 2

**Réponses exactes : B – C – E**

**Inadmissible : A**

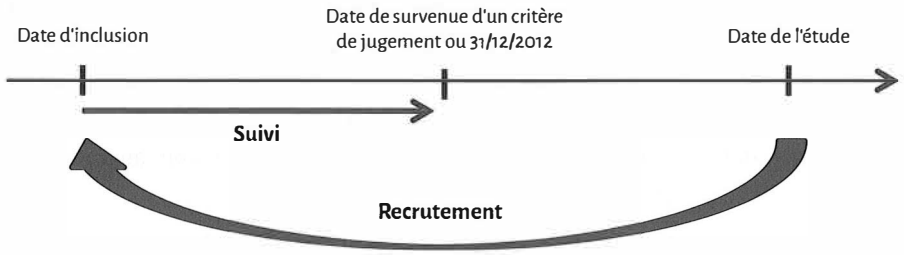
*Commentaires* Les investigateurs de cette étude ne maîtrisaient pas l'attribution de l'exposition ici l'évaluation par coroscanner (notamment par randomisation). Ce choix était laissé à l'équipe médicale qui prenait en charge le patient. De ce fait cette étude n'est pas expérimentale mais observationnelle. Il s'agit d'une étude de cohorte (« prospective cohort of ischemic stroke patients » en début de partie Materials and Methods) incluant des patients exposés (c'est-à-dire ayant bénéficié de l'évaluation par coroscanner) et de patients non-exposés (c'est-à-dire n'ayant pas bénéficié de l'évaluation par coroscanner).

### Question 3

**Réponse exacte : B**

**Inadmissible : D**

*Commentaires* Les auteurs qualifient cette étude de « retrospective study of a prospective cohort ». Ici il faut comprendre que les données ont été consignées dans les dossiers médicaux (caractéristiques cliniques à l'inclusion, évaluation par coroscanner, et critère de jugement) lors du suivi des patients. Le suivi des sujets était donc effectué de manière prospective puisqu'il correspondait au suivi médical des patients. Par contre la décision d'inclure ces patients dans une étude a été prise ultérieurement au suivi médical de ces patients et le recrutement était rétrospectif. On peut schématiser ainsi le déroulement :




### Question 4

**Réponse exacte : D**

*Commentaires* Cette étude s'est intéressée à la survenue de 3 types d'événements distincts et considérés séparément : le décès, la première occurrence d'un événement cardiovasculaire ou d'une récidive d'AVC. « Outcomes measured in this study were death and first occurrence



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

of cardiovascular events or recurrent ischemic stroke » (partie Materials and Method section Outcome data). Dans une analyse de survie comme effectuée dans cette étude, le critère de jugement est le délai de survenue de l'événement. Ce sont des délais jusqu'à la survenue du décès, d'un événement cardiovasculaire ou d'une récidive d'AVC qui ont constitué les 3 critères de jugement principaux (« We obtained data for vascular events and death (date and causes of death) by review of the medical records and face-to-face or telephone interviews carried out by the stroke research nurse and stroke specialists. »). Pour certains sujets, les événements considérés ne se sont pas produits durant le suivi, le suivi a été censuré à la date de point (date de fin d'étude), le 31/12/2012. Aucun critère de jugement n'a combiné les 3 événements. Les auteurs ont réalisé les analyses pour les 3 délais de façon distincte. Les codes diagnostic de la CIM 10 étaient utilisés pour identifier les événements cardiovasculaires et récidives d'AVC et également pour reclasser les décès en événements cardiovasculaires ou récidives d'AVC.

■ **Question 5**

**Réponse exacte : C**

**Indispensable : C**

*Commentaires* L'objectif de cette étude concernant l'effet de l'évaluation par coroscanner sur la survenue à long terme (selon les auteurs) de 3 types d'événements : décès, événement cardio-vasculaire ou cérébrovasculaire. On retrouve l'objectif en fin d'introduction : « we investigated whether long-term cardio-cerebrovascular outcomes differ between patients who underwent MDCT during the acute stage of stroke and those who did not. » Les auteurs considèrent un effet à long terme, ce qui peut être discutable au regard du suivi moyen qui était de 38 mois. Il faut noter que les critères d'exclusion prévoyaient d'exclure les patients avec des antécédents de coronaropathie.

■ **Question 6**

**Réponses exactes : B – C**

**Inadmissibles : D – E**

*Commentaires* L'évaluation par coroscanner avait pour but d'identifier des coronaropathies asymptomatiques (« For determination of the presence of asymptomatic CAOD in stroke patients, since July 2006, MDCT was consecutively performed » en partie Materials and Methods section Study design and population 2e paragraphe). Exclure les sujets prévalents (c'est-à-dire pour lesquels une coronaropathie était déjà connue) était nécessaire, avoir une coronaropathie était d'ailleurs un critère d'exclusion du coroscanner chez ces sujets. L'exclusion des sujets qui décédaient dans les 7 séjours après l'AVC s'expliquait en raison d'un manque de temps pour réaliser le coroscanner. Si ces sujets n'avaient pas été exclus, un biais de survie aurait pu apparaître car les sujets décédés précocement (notamment avant d'avoir pu bénéficier du coroscanner) seraient classés dans le groupe n'ayant pas bénéficié du coroscanner et donc augmenterait le taux de décès de ce groupe artificiellement.

Cette étude étant monocentrique (« the neurology department of the Severance Stroke Center, Yonsei University in Seoul, Korea ») la proposition D est fausse. Les auteurs justifient l'absence de consentement éclairé par le caractère rétrospectif de l'étude.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 7

**Réponses exactes : C – D**

**Indispensable : C**

*Commentaires* Le tableau 1 montre que les sujets ayant bénéficié de l'évaluation par coroscanner et ceux n'en ayant pas bénéficié avaient des caractéristiques différentes notamment en termes d'âge (respectivement 64 ans vs 68 ans) et de sexe (respectivement 64 % d'hommes vs 58 %). Le texte (partie Results section Baseline characteristics of the MDCT (+) and MDCT (-) groups 2e paragraphe) mentionne que 62,7 % des sujets ayant bénéficié de l'évaluation par coroscanner avaient une coronaropathie asymptomatique, chez les sujets n'ayant pas bénéficié du coroscanner, il est impossible de connaître ce pourcentage. La proposition A est donc fausse. Dans le tableau 1, on peut lire que la proportion de sujets diabétiques dans les groupes ayant bénéficié du coroscanner est de 30,7 % contre 33,7 % dans le groupe n'en ayant pas bénéficié. La proportion de sujets diabétiques dans le groupe ayant bénéficié du coroscanner était donc bien supérieure. Le résultat du test statistique en regard est non significatif (p = 0,072). Cela signifie que la différence observée entre ces 2 proportions n'existe pas au niveau des populations dont sont extraits ces 2 échantillons. La proportion de diabétiques n'est donc pas différente entre la population de l'ensemble des sujets bénéficiant du coroscanner (de laquelle sont extraits les 1842 sujets inclus) et la population de l'ensemble ne bénéficiant pas du coroscanner (de laquelle sont extraits les 1275 sujets inclus). Dans le texte (partie Results section Cardiovascular and cerebrovascular outcomes paragraphe Outcomes in the original population), les auteurs donnent une durée de suivi moyenne de 38 mois avec une valeur minimale de 0 mois et maximale de 85,9 mois (« During a follow-up of 38,0±24,8 months (range : 0–85,9 months) ». En fonction de la date d'inclusion, de la survenue ou non de l'un des 3 types d'événement considéré, un sujet avait un suivi plus ou moins long. Par exemple un sujet ayant une récidive d'AVC 10 mois après son inclusion et avant le 31/12/2012 avait un suivi dans l'étude de 10 mois, alors qu'un sujet décédant 2 jours après son inclusion et avant le 31/12/2012 avait un suivi dans l'étude de 2 jours. Dans la suite du paragraphe, les auteurs détaillent le nombre de morts : 482, le nombre d'événements cardiovasculaires : 60 et le nombre de récidives d'AVC : 297.

## Question 8

**Réponses exactes : C – D – E**

*Commentaires* Dans le tableau 1, on constate que 2 échantillons d'analyse ont été considérés : l'échantillon complet avec les 3 117 sujets et l'échantillon apparié sur le score de propension avec les 1616 sujets : 808 sujets ayant bénéficié du coroscanner et 808 sujets n'en ayant pas bénéficié. Le tableau 1 montre également que les caractéristiques initiales des sujets inclus dans cette analyse sont proches sans toutefois être identiques (par exemple sur la proportion de sujets diabétiques 76,7 % pour le groupe ayant été évalué par coroscanner contre 77,1 % dans le groupe n'en ayant pas bénéficié). Les résultats des tests statistiques effectués sur les caractéristiques présentées ne permettent pas de rejeter l'hypothèse que les 2 échantillons de 808 sujets issus d'une même population. L'appariement sur le score de propension permet de sélectionner des échantillons de sujets dans chaque groupe qui globalement ont des caractéristiques proches. Il est cohérent d'observer que les caractéristiques de ces sujets sélectionnés sont situées entre celles des 2 groupes non appariés. Toutefois il peut arriver que ce ne soit pas le cas.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 9**

**Réponse exacte : A**

*Commentaires* La figure 2 permet d'affirmer que les sujets de l'étude ayant bénéficié d'une évaluation par coroscanner avaient un meilleur taux de survie (figure 2A), présentaient moins d'événements cardiovasculaires (figure 2B) et moins de de récidives d'AVC (figure 2C). La mention « survival probability » sur chaque figure fait référence au type d'analyse (analyse de survie) mais l'événement considéré n'est pas forcément le décès ; pour les figures B et C les événements sont respectivement les événements cardiovasculaires et la récidive d'AVC. La figure 2 ne permet pas de conclure à l'effet de l'évaluation par coroscanner sur les populations en raison d'un biais majeur de confusion (lié à l'absence de randomisation) à savoir que les caractéristiques initiales des populations dont sont issus les 2 échantillons (celui de 1842 pour ceux ayant bénéficié du coroscanner et de 1275 pour ceux n'en ayant pas bénéficié) ne sont pas comparables ($p < 0,05$ pour de nombreuses caractéristiques dans le tableau 1).

**Question 10**

**Réponses exactes : A – E**

**Indispensable : A**

*Commentaires* Le score de propension correspond à la probabilité pour un sujet de bénéficier de l'exposition (ici l'évaluation par coroscanner) à partir de ses caractéristiques initiales (partie Materiels and methods section Statistical analysis « Each patient was then assigned an estimated propensity score, which was the predicted probability of undergoing MDCT on the basis of patient's observed baseline characteristics »). Comme toute probabilité elle varie entre 0 (impossibilité pour le sujet d'être exposé) et 1 (certitude pour le sujet d'être exposé). Elle est calculée pour tous les sujets qu'ils aient ou non bénéficié de l'évaluation par coroscanner à partir de leurs caractéristiques initiales (sexe, âge, antécédent de diabète, d'hypertension...) détaillés dans le paragraphe. Ce calcul a été obtenu à partir d'une régression logistique modélisant le fait d'avoir bénéficié du coroscanner (et non pas la survie), modèle pour une variable à expliquer binaire et non modèle de durée jusqu'à un événement comme le modèle de Cox. Les sujets ayant bénéficié du coroscanner ont ensuite été appariés à ceux n'en ayant pas bénéficié sur ce score de propension, ces paires avaient donc un score de propension identique (ou très proche).

**Question 11**

**Réponses exactes : A – D**

**Inadmissible : E**

*Commentaires* Le tableau 2 présente pour chacun des 3 critères de jugement 7 types de modèles dont 2 analyses classiques : analyse de survie non ajustée, analyse de survie ajustée, et 5 analyses basées sur le score de propension : analyse de survie stratifiée sur le score de propension, analyse de survie dans chacun des quintiles du score de propension, analyse de survie ajustée sur le score de propension, analyse de survie pondérée par l'inverse du score de propension et analyse de survie appariée sur le score de propension. L'analyse dans chaque quintile du score de propension avait une puissance réduite par rapport aux autres analyses multivariées en raison d'un effectif (et donc d'un nombre d'événements) diminué dans chacune des strates.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 12

**Réponses exactes : C – D**

*Commentaires* Dans l'analyse de survie dans chaque quintile du score de propension du tableau 2, le premier quintile correspond aux sujets ayant la plus faible probabilité d'avoir été évalués par coroscanner et le 5e quintile les sujets ayant la plus forte probabilité d'avoir été évalués par coroscanner. Dans le tableau 2, toutes les analyses modélisent le fait de ne pas être évalué par coroscanner qui est donc considéré comme la référence HR = 1 (et non pas le fait d'être dans le premier quintile du score de propension). Les sujets appartenant au 1er quintile n'avaient aucun effet significatif de l'évaluation par coroscanner que ce soit sur la mortalité, les événements cardiovasculaires ou la récidive d'AVC ($p > 0,05$). Les sujets du 5e quintile étaient les seuls à avoir un effet néfaste significatif (HR > 1 et $p < 0,05$) de l'absence d'évaluation par coroscanner (donc un effet bénéfique de l'évaluation par coroscanner) sur les événements cardiovasculaires et la récidive d'AVC.

### Question 13

**Réponses exactes : A – B – C**

*Commentaires* Dans le texte (partie Results section Outcomes in the propensity score-matched population « In addition, we performed comparative analyses using the data of the MDCT (+) group (n = 1842) and 606 patients of MDCT (-) subgroup who had not undergone MDCT due to non-medical reasons », les auteurs mentionnent cette analyse de sensibilité qui n'a retrouvé que l'association significative entre le décès et l'évaluation par coroscanner dans les analyses basées sur le score de propension.

### Question 14

**Réponses exactes : A – B – C – E**

**Inadmissible : D**

*Commentaires* Les limites de l'étude sont discutées par les auteurs dans la partie Discussion (avant dernier paragraphe). Ils évoquent l'absence de randomisation, la possibilité de facteurs de confusion non pris ne compte tel que le statut socio-économique, le caractère monocentrique. Ce dernier critère pourrait limiter la généralisation des résultats de l'étude par un biais de sélection si ce centre a recruté des patients ayant une particularité.

### Question 15

**Réponses exactes : A – E**

*Commentaires* Le financement de l'étude est mentionné dans la partie Acknowledgements. Les affiliations des auteurs permettent d'identifier leurs structures d'appartenance (université et hôpital). Sur la première page les auteurs déclarent ne pas avoir de conflit d'intérêts.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 3 Étude de cohorte

Pascal Auquier, Julie Berbis

| | |
|---|---|
| **Titre de l'article** | **Risk of cancer in patients with polycystic kidney disease: a propensity-score matched analysis of a nationwide, population-based** |
| **Auteurs** | Yu TM, Chuang YW, Yu MC, Chen CH, Yang CK, Huang ST., Lin CL., Shu KH., Kao CH |
| **Référence** | Lancet Oncol. 2016 Oct ; 17(10):1419-1425. |
| **Lien vers l'article** | **doi.org/10.1016/S1470-2045(16)30250-9** |

**1** **Quel(s) est (sont) le(s) objectif(s) annoncé(s) de l'étude ?**

a. Étudier l'incidence de la polykystose rénale chez l'adulte
b. Étudier le lien entre maladie polykystique rénale et cancer
c. Étudier le risque de survenue de cancer chez les patients en insuffisance rénale chronique
d. Étudier la prévalence des cancers chez les patients présentant une polykystose
e. Étudier le rôle du cancer dans la survenue de la maladie polykystique

**2** **Quel(s) est (sont) le(s) argument(s) rapportés par l'auteur justifiant la mise en place de l'étude ?**

a. La maladie polykystique du rein est une affection pouvant avoir une évolution sévère
b. Il existe un modèle physio pathologique
c. Les gènes impliqués dans la polykystose rénale sont exprimés dans de nombreux tissus
d. Les études préalables, montrant un lien entre polykystose rénale et cancer, sont de faibles qualités
e. La maladie polykystique du rein est une affection rare

**3** **Concernant l'objectif de l'étude, parmi les propositions suivantes laquelle ou lesquelles sont vraie(s) ?**

a. Il est défini clairement
b. Il est justifié par les éléments de la conclusion
c. Il justifie la mise en place d'une étude expérimentale
d. Il justifie la mise en place d'une étude descriptive
e. Il est d'identifier un facteur potentiel de risque



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **Concernant la population cible de l'étude, parmi les propositions suivantes laquelle ou lesquelles sont vraie(s) ?**

a. La population cible est l'ensemble des patients présentant une polykystose rénale sans insuffisance rénale chronique
b. La population cible est l'ensemble des patients présentant une polykystose rénale à Taiwan
c. La population cible est l'ensemble des patients présentant une polykystose rénale
d. La population cible est l'ensemble des patients adultes présentant une polykystose rénale
e. La population cible est l'ensemble des patients présentant un cancer

**5** **Parmi les propositions suivantes laquelle ou lesquelles sont vraie(s) ?**

a. L'étude est une étude expérimentale
b. L'étude est descriptive
c. L'étude est analytique
d. L'étude est de type « exposé » non exposé
e. L'étude est de type cas témoins

**6** **Parmi les propositions suivantes laquelle ou lesquelles sont vraie(s) ?**

a. Le choix de témoins non exposés apparentés aurait été moins biaisé
b. L'absence d'information sur le tabagisme constitue un biais
c. L'absence d'information sur le BMI, comme traceur des conduites alimentaires et de l'activité physique constitue un biais
d. La codification proposée des cancers est inadaptée
e. Le choix des patients présentant une polykystose rénale dans la base nationale de l'assurance maladie constitue un biais

**7** **Concernant l'analyse prévue, parmi les propositions suivantes laquelle ou lesquelles sont vraie(s) ?**

a. L'analyse prend en compte les facteurs de confusion
b. Le score de propension permet un appariement d'un sujet non exposé plus semblable au patient exposé (polykystose rénale)
c. L'appariement prend en compte le tabagisme
d. La prise en compte de la survenue du décès dans l'analyse semble justifiée
e. L'analyse aurait dû prendre en compte les comorbidités après la date index afin d'assurer la comparabilité des groupes

**8** **Concernant le tableau 1, parmi les propositions suivantes laquelle ou lesquelles sont vraie(s) ?**

a. L'âge médian diffère significativement entre les sujets exposés et non exposé
b. 75 % de la population non exposé a plus de 73,1 ans
c. 25 % de la population exposé a plus de 73,1 ans
d. 25 % de la population non exposé a moins de 44,4 ans
e. 95 % de la population non exposé a plus de 73,1 ans



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**9** **Concernant les résultats, parmi les propositions suivantes laquelle ou lesquelles sont vraie(s) ?**

a. 26 % des sujets hypertendus présentent une polykystose rénale
b. L'égalité des effectifs entre les deux groupes exposés non exposés constitue un biais
c. Le nombre de cancer chez les sujets présentant une polykystose rénale dans la population cible est de 20,1
d. Le nombre de cancer chez les sujets ne présentant pas de polykystose rénale dans la population cible peut être 18,3
e. Le nombre de cancer est plus important chez les sujets présentant une polykystose rénale

**10** **Concernant les résultats rapportés au tableau 2, parmi les propositions suivantes laquelle ou lesquelles sont vraie(s) ?**

a. Le risque brut lié à la présence d'une polykystose rénale est de 1,83
b. Le modèle prend en compte l'ensemble des facteurs de confusion définis
c. Le risque prenant en compte les facteurs de confusion est significativement plus élevé chez les hommes que chez les femmes
d. Le risque prenant en compte les facteurs de confusion est significativement plus élevé chez les sujets présentant une polykystose rénale comparativement aux sujets indemnes
e. Le tableau prend en compte le risque compétitif de décès

**11** **Dans cette étude, le risque de conclure à tort à la responsabilité de la polykystose rénale dans la survenue d'un cancer peut s'expliquer par**

a. Le risque d'erreur statistique de première espèce (alpha)
b. Un défaut de puissance statistique
c. L'existence de facteurs de confusion résiduels méconnus
d. L'existence d'un risque compétitif lié au décès
e. Le caractère rétrospectif de l'étude

**12** **Au regard du tableau 3, parmi les propositions suivantes laquelle ou lesquelles sont vraie(s) ?**

a. Dans la population étudiée, la maladie polykystique du rein diminue le risque de cancer de la thyroïde
b. Le lien significatif avec la survenue du cancer du pancréas est probablement un surrisque lié à une autre cause non prise en compte
c. Au regard des risques ajustés, la population cible a au moins deux fois plus de risque de faire un cancer du rein qu'un cancer du poumon
d. Le risque ajusté de cancer du pancréas étant plus élevé que celui du cancer du côlon, le nombre de cancer du pancréas sera donc plus élevé
e. L'objectif étant centré sur la survenue du cancer, les auteurs ne peuvent rapporter les résultats par type de cancer



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**13** **Parmi les propositions suivantes, concernant les résultats rapportés par les auteurs, laquelle ou lesquelles sont vraie(s) ?**

a. Au regard du tableau 2 la maladie polykystique du rein est un facteur de risque de la survenue du cancer
b. La relation dose-effet est un argument en faveur de la causalité
c. La répartition relativement proche des événements de comorbidités dans les deux groupes (exposés et non exposés) remet en cause la qualité du groupe comparateur
d. L'absence d'association similaire à d'autres facteurs que la maladie polykystique rénale et le cancer serait un argument de causalité
e. L'antériorité de l'exposition n'est pas un argument causal

**14** **Concernant la discussion, parmi les propositions suivantes laquelle ou lesquelles sont vraie(s) ?**

a. L'opinion exprimée par les auteurs dans la première phrase de la discussion est juste
b. L'opinion exprimée par les auteurs dans la dernière phrase du premier paragraphe de la discussion est juste
c. Il existe des études confirmant le résultat trouvé dans certains cancers
d. Cette étude permet de faire l'hypothèse d'un lien entre maladie polykystique comme précurseur de maladie cancéreuse
e. Les études rapportées en discussion montrant des liens similaires sont un argument en faveur de la causalité

**15** **Parmi les propositions suivantes laquelle ou lesquelles sont vraie(s) ?**

a. Pour répondre à l'objectif posé par les auteurs, il n'était pas possible de mettre en place une étude prétendant à un niveau de preuve 1 (selon la Haute Autorité de Santé)
b. Les résultats sont extrapolables à la population cible
c. Au regard du risque augmenté de cancer du côlon rapporté dans l'étude vous recommanderiez à vos patients présentant une polykystose rénale un dépistage systématique
d. Au regard du risque non significatif de survenue de cancer du sein vous recommanderiez à vos patients présentant une polykystose rénale de ne pas participer aux campagnes de dépistage collectif
e. Au regard des résultats, vous ne lirez pas les résultats d'une prochaine étude sur le même sujet



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 1

**Réponse exacte : B**

*Commentaires*

La réponse A est fausse : l'étude est une étude explicative et non descriptive (Étudier l'incidence de la polykystose rénale chez l'adulte).

La réponse B est juste : cette réponse correspond à l'énoncé de l'objectif en fin d'introduction (« we did a nationwide cohort study to establish the risk of cancer in patients with polycystic kidney disease without chronic kidney disease or end-stage renal disease ») qui est cohérent avec le titre (« Risk of cancer in patients with polycystic kidney disease : a propensity-score matched analysis of a nationwide, population-based cohort study »).

La réponse C est fausse : le facteur suspecté est la polykystose rénale à l'exclusion de l'insuffisance rénale chronique (introduction : « without chronic kidney disease or end-stage renal disease ») qui n'est pas prise en considération dans l'étude. La réponse D est fausse : l'étude est une étude explicative et non descriptive (Étudier la prévalence).

La réponse E est fausse : la polykystose rénale est suspectée d'être un facteur explicatif du cancer et non l'inverse.

## Question 2

**Réponses exactes : A – B – C**

*Commentaires*

La réponse A est juste : la maladie polykystique rénale peut évoluer vers des anévrysmes des artères cérébrales mais aussi d'autres pathologies (introduction : « colon diverticular disease, and liver cysts »).

La réponse B est juste : le modèle physio pathologique et les mécanismes en jeu sont rapportés en introduction avec des éléments référencés et repris dans la discussion.

La réponse C est juste : les auteurs rapportent « these genes are widely expressed in several tissues throughout the body ».

La réponse D est fausse : aucune étude explorant le lien maladie polykystique du rein et cancer n'est rapportée par les auteurs en introduction (cependant certaines sont citées en discussion sur le cancer du rein).

La réponse E est fausse : la maladie polykystique du rein est une maladie fréquente ; la rareté de l'exposition (ici la maladie polykystique rénale) est un argument pour la mise en place d'une étude exposé non exposé contre un choix d'étude cas-témoins.

## Question 3

**Réponses exactes : A – E**

*Commentaires*

La réponse A est juste : l'objectif défini en fin d'introduction est cohérent avec le titre de l'article ; il apparaît clairement que l'objectif est d'étudier le rôle de la maladie polykystique rénale comme potentiel facteur de risque (mise en évidence d'un lien) dans la survenue de cancers.

La réponse B est fausse : la justification d'une étude repose sur les éléments rapportés dans l'introduction, parfois dans le chapitre discussion ce qui n'est pas conventionnel.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

La réponse C est fausse : l'objectif repose sur la mise en évidence d'un lien, la variable étudiée (explicative) est la maladie polykystique (la variable à expliquer est le cancer). La variable explicative n'est pas maitrisable par l'investigateur principal : il ne peut la randomiser. Le schéma d'étude repose donc sur une étude étiologique (ou analytique) et non un schéma expérimental randomisé.

La réponse D est fausse : l'objectif de l'étude est explicatif, l'étude ne peut donc être descriptive qui ne permet pas d'interprétation causale.

La réponse E est juste : la maladie polykystique du rein est identifiée par les auteurs comme un facteur potentiel de risque.

### Question 4

**Réponse exacte : C**

*Commentaires*

La réponse A est fausse : la population cible est celle sur laquelle on (l'auteur) souhaite faire porter les résultats : extrapolation. Elle est définie dans le titre et est conforme à celle de la conclusion à savoir ici l'ensemble des patients présentant une polykystose rénale. La confusion vient qu'ici la population qui est rapporté en fin d'introduction dans l'objectif est restreinte à un sous-ensemble plus proche de la population source.

La réponse B est fausse : la population cible est l'ensemble des patients présentant une polykystose rénale. Au chapitre matériel, cette population est prise dans une base de données à Taiwan qui est la population source. On peut se questionner sur la capacité de cette population source à recouvrir la cible (extrapolation).

La réponse C est juste : la population cible est celle sur laquelle on (l'auteur) souhaite faire porter les résultats : extrapolation. Elle est définie dans le titre et est conforme à celle de la conclusion à savoir ensemble des patients présentant une polykystose rénale.

La réponse D est fausse : au chapitre matériel on comprend que la population incluse ne concerne que les adultes à partir de 20 ans : population source.

La réponse E est fausse : le cancer est le phénomène morbide étudié c'est-à-dire la variable à expliquer qui ne peut être à l'origine de la définition de la population cible.

### Question 5

**Réponses exactes : C – D**

*Commentaires*

La réponse A est fausse : une étude expérimentale implique que le facteur explicatif étudié soit attribué de façon aléatoire entre les groupes comparés. Les groupes comparés sont : patients avec polkystose rénale *versus* sans polykystose rénale, la polykystose n'est pas « attribuée » par randomisation elle est un fait constaté.

La réponse B est fausse : l'objectif de l'étude est rechercher un lien potentiel entre polykystose rénale et cancer : l'étude est donc explicative et non descriptive.

La réponse C est juste : l'étude est explicative. La variable explicative (polykystose) n'est pas attribuable au hasard (randomisable). L'étude est donc analytique ou étiologique.

La réponse D est juste : l'étude est analytique (ou étiologique). Elle peut être soit de type exposé-non exposé soit cas-témoins. Ici l'auteur prend un groupe de personnes exposé (polykystose rénale) un groupe non exposé indemnes de polykystose rénale) et on observe de façon prospective dans les 2 groupes la survenue de cancer (variable à expliquer).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

La réponse E est fausse : l'étude serait de type cas (cancer) témoins (non-cancer) si un groupe témoin des cancers était considéré (ce qui n'est pas le cas) et que l'on recherchait antérieurement l'existence d'une polykystose rénale dans ces deux groupes (ce qui n'est pas le cas).

### Question 6

**Réponses exactes : B – C**

*Commentaires*

La réponse A est fausse : le choix de témoins apparentés, qui peut être justifié dans certaines études aurait pu créer le risque de sur-appariement, masquant le risque lié à la polykystose rénale. Par ailleurs les auteurs ont la possibilité de choisir les témoins non exposés dans la base des patients bénéficiant de l'assurance maladie dont on peut penser qu'elle est représentative de la population ; sujets exposés et non exposés issus de la même source.

La réponse B est juste : les auteurs étudient la survenue d'un très grand nombre de cancers. Il est donc difficile de disposer du relevé de l'ensemble des facteurs de risques connus et spécifiques de chaque type de cancer. Cependant, le tabagisme, le surpoids et dans une moindre mesure la consommation d'alcool, l'exposition aux UV constituent des facteurs de risque retrouvés dans de nombreux cancers. Le tabagisme pourrait être inégalement réparti entre les sujets exposés et non exposé, entraînant une sur ou sous-estimation du risque. On parle alors de facteur de confusion qui est facteur lié à la variable explicative et à la variable à expliquer.

La réponse C est juste : *cf.* la correction de la réponse B.

La réponse D est fausse : les auteurs recourent à la classification internationale des maladies (ICD) qui est une classification consensuelle et partagée permettant de comparer des résultats d'études sur le même sujet.

La réponse E est fausse : les auteurs disposent d'une base nationale des patients. Cette base permet une collection exhaustive de la pathologie sur le territoire national, ce qui limite les biais d'extrapolation. En général l'enregistrement de « cas » dans les bases de l'assurance maladie répond à des critères diagnostiques (cliniques, paracliniques) précis et validés. Sous réserve que la majorité de la population bénéficie de l'assurance maladie (et l'auteur indique dans la méthode 99 % de la population), alors le choix de recourir à cette base pour répondre à l'objectif semble la meilleure des stratégies (biais faible).

### Question 7

**Réponses exactes : A – B – D**

*Commentaires*

La réponse A est juste : le score de propension permet de définir un sujet comparateur (non exposé) le plus proche possible du sujet exposé en prenant en compte les principaux facteurs de confusion.

La réponse B est juste : voire justification ci-dessus.

La réponse C est fausse : le tabagisme est un facteur de risque connu de la survenue d'un large nombre de cancers. Il n'est pas ici pris en compte dans l'analyse. Les auteurs au chapitre discussion soulignent cette limite.

La réponse D est juste : la survenue d'un décès après la constitution des 2 groupes exposé et non exposé, empêche la survenue et la comptabilisation d'un cas de cancer potentiel (puisqu'un sujet décédé n'a plus la possibilité de développer un cancer, c.-à-d. l'évènement surveillé, par la suite). Le décès peut donc surexprimer l'importance d'un lien, il constitue donc un risque



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

qui est en compétition avec la variable à expliquer (le cancer) on parle de risque compétitif et il est normal de prendre en compte cette variable dans l'analyse comme potentiel facteur de confusion.

La réponse E est fausse : la prise en compte des comorbidités comme facteur de confusion permet d'assurer la comparabilité des groupes à l'inclusion. Le recueil de l'information doit ainsi concerné l'information avant la date index qui est la date du diagnostic.

## Question 8

**Réponse exacte : D**

*Commentaires*

La réponse A est fausse : les âges moyens sont respectivement de 57,6 (non exposé) et 60,0 (exposé), et les auteurs précisent dans le texte qu'ils n'ont pas mis en évidence de différence significative sur les principales caractéristiques excepté sur l'hypertension et le diabète. La troisième colonne rapporte la différence des moyennes standardisée, ici de 0,034, qu'il ne faut pas confondre avec probabilité d'erreur (p. value).

La réponse B est fausse : l'intervalle interquartile rapporté pour l'âge de la population non exposée est de [44,4 – 73,1], ce qui veut dire que 25 % de cette population a un âge de moins de 44,4 ans et 25 % a un âge de plus de 73,1 ans.

La réponse C est fausse : 25 % de la population non exposé a plus de 73,1 ans mais la question porte sur la population exposée dont le quartile supérieur est de 73,8 et non de 73,1.

La réponse D est vraie : *cf.* la réponse faite à la question B.

La réponse E est fausse : *cf.* réponse B, la confusion serait de confondre l'intervalle interquartile et l'intervalle de confiance à 95 %.

## Question 9

**Réponse exacte : E**

*Commentaires*

La réponse A est fausse : au tableau 1 1106 sujets parmi les 4346 patients présentant une polykystose rénale présentent une hypertension soit 26 % ; à l'inverse l'étude ne permet de connaître dans la population hypertendue le nombre de patients présentant une polykystose.

La réponse B est fausse : les auteurs annoncent dans le chapitre méthodes (analyse statistiques) que chaque sujet exposé été apparié avec un sujet non exposé sur la base d'un score de propension qui permet de choisir le sujet le plus « proche » au regard des facteurs de confusion considérés. Cette approche a pour but de limiter les biais. Il est donc rassurant d'observer au chapitre résultat que les tailles des populations exposé et non exposé sont identiques.

La réponse C est fausse : le taux d'incidence est de 20,1 dans la population étudiée ; classiquement le résultat est fourni sur une population de résultat issue d'une source. L'intervalle de confiance est celui pour la population source. Ici la population de résultat est la même que la source, l'intervalle de confiance concerne donc la cible. Dans la population cible il y a 95 chances sur 100 que la vraie valeur de ce taux soit compris entre 18,3 et 21,9 (intervalle de confiance) soit une quelconque de ces valeurs dont 20,1, mais il n'y a pas 95 chances sur 100 que cette valeur soit 20,1.

La réponse D est fausse : au regard de la réponse précédente cette proposition est juste pour la population exposé (polykystose) mais fausse pour les non exposés (IC = [10,1-11,8]).

La réponse E est juste : le nombre est celui observé dans les résultats soit 404 contre 261.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 10

**Réponses exactes : B – D – E**

*Commentaires*

La réponse A est fausse : 1,83 est le risque ajusté après prise en compte des facteurs de confusion et non le risque brut non ajusté.

La réponse B est juste : au chapitre méthodes les auteurs annoncent vouloir prendre en compte l'âge, le genre, la fréquence des consultations et les comorbidités ; le tableau de résultats rapporte un Hasard-Ratio ajusté sur ces 4 facteurs (commentaires en bas du tableau).

La réponse C est fausse : les risques ajustés sont respectivement chez les femmes et les hommes de 1,65 et 1,93. Ces deux risques sont significatifs (lecture de l'intervalle de confiance ou de la p value) dans chacune des strates : on observe un lien entre polykystose rénale et cancer dans la strate des hommes ; on observe aussi ce lien dans la strate des femmes. Cependant la différence observée homme *versus* femme n'est pas comparée. Les risques ajustés observés diffèrent mais on ne peut conclure que cette différence n'est pas l'effet du hasard. L'effet du genre n'est pas testé.

La réponse D est juste : le risque prenant en compte les facteurs de confusion est le HR ajusté. Il y a 95 % de chances que la vraie valeur du risque prenne n'importe quelle valeur entre 1,57 et 2,15. Ce risque est donc significatif ce que confirme le test de signification (p. value).

La réponse E est juste : les auteurs comme prévu au chapitre méthode fournissent en complément du risque ajusté, un risque ajusté prenant en compte le décès (Adjusted SHR).

## Question 11

**Réponses exactes : A – C**

*Commentaires*

La réponse A est juste : le risque de 1re espèce alpha est le risque de conclure à une différence dans la population cible (classiquement avec 5 %) alors que cette différence n'existe pas.

La réponse B est fausse : le défaut de puissance est lié au risque bêta qui est le risque de conclure à l'absence de différence dans la population cible (au risque bêta par exemple 20 %) alors que cette différence existe ; une étude ayant un risque bêta de 20 % aura 80 % de chance de mettre en évidence une différence (puissance) si cette différence existe vraiment.

La réponse C est vraie : des facteurs de confusion connus ou non connus peuvent être répartis différemment entre les deux groupes ; dans le cas présent le tabagisme est un facteur de confusion important qui n'a pas pu être pris en compte car probablement non renseigné dans les bases de l'assurance maladie.

La réponse D est fausse : le décès est un risque compétitif pertinent et argumenté par les auteurs qui rapportent les résultats ajustés avec et sans prise en compte de ce risque. Ce facteur n'est donc pas ignoré et par ailleurs les deux risques bruts (et leurs intervalles de confiance) semblent proches.

La réponse E est fausse : l'étude est constituée rétrospectivement à un instant t sur une base de données préexistante dans laquelle le recueil des données a été prospectif.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 12

**Réponses exactes : A – B**

*Commentaires*

La réponse A est vraie : le risque qu'il soit brut ajusté avec ou sans risque compétitif est toujours inférieur à 1 (de 0,67 à 0,74) ; les patients de l'étude présentant une polyskystose rénale ont développé moins de cancers de la thyroïde.

La réponse B est vraie : le lien retrouvé est fort 5,37 et significatif, cependant la prise compte du risque compétitif lié au décès montre que ce lien diminue 1,88 et qu'il est non significatif. Il est probable que ces sujets décédés soient des cancers liés à l'alcool, facteur non pris en compte dans l'analyse.

La réponse C est fausse : les risques ajustés sont de 5,26 et1,99 (soit un rapport de 2,6) ; l'examen des intervalles de confiance montrent que si ces deux risques sont significatifs cependant les valeurs possibles dans la population cible font que ce rapport ne sera pas systématiquement multiplié par plus de 2 (par exemple si la vraie valeur est dans les deux cas la borne inférieure de l'intervalle de confiance).

La réponse D est fausse : si le risque rapporté est plus élevé dans le cancer du pancréas comparativement au colon, le risque reste une valeur relative qui n'est pas une quantification et le dénombrement montre un nombre plus élevé (comme attendu) de cancer du côlon (qui est fréquent) que de cancer du pancréas (qui est plus rare).

La réponse E est fausse : ce qui n'est plus maitrisé ce sont les probabilités d'erreur données pour les risques associés à chaque cancer ; cependant une étude exposée non-exposé permet de rapporter plusieurs événements ; le résultat par type de cancer est alors documentaire.

## Question 13

**Réponse exacte : D**

*Commentaires*

La réponse A est fausse : le tableau 2 met en évidence un risque significatif. La significativité d'un résultat dans une étude étiologique n'implique pas la causalité du fait de la non maitrise de la comparabilité des groupes ; par construction les études étiologiques ont des niveaux de preuve moindre que les études expérimentales randomisées. L'interprétation causale se fait au regard des éléments propre à l'étude et d'éléments extérieurs.

La réponse B est fausse : en général dans les études étiologiques la relation dose effet est un élément de la discussion causale. Ici cette relation n'est pas documentée, la réponse est donc fausse.

La réponse C est fausse : le groupe non exposé n'est pas synonyme de groupe de sujets sains ; il est constitué de sujet non soumis à l'exposition ici la maladie (polykystose rénale) ; par ailleurs les auteurs rapportent qu'ils ont réalisés un appariement (1 :1) probabilistique par score de propension prenant en compte plusieurs facteurs dont l'un est la comorbidité, il est donc attendu que les nombres d'événements rapportés à la durée d'exposition (personne-année) soient proches entre les deux groupes.

La réponse D est vraie : la spécificité de l'effet est un argument en faveur de la causalité dans les études étiologiques ; cette spécificité n'est pas documentée ici mais la question est posée en général.

La réponse E est fausse : pour pouvoir entreprendre une discussion causale il faut s'assurer de la préexistence de l'exposition (ici la maladie polykystique rénale) à toute survenue de l'événement (ici le cancer).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 14

**Réponses exactes : A – B – C – D**

*Commentaires*

La réponse A est vraie : dans la cohorte (population de résultat) le risque est augmenté.

La réponse B est vraie : les auteurs évoquent un lien ce qui est le cas.

La réponse C est vraie : les auteurs rapportent les résultats d'étude en faveur du lien maladie polykystique et cancer du rein dans la discussion.

La réponse D est vraie : l'étude ne permet pas en l'état de conclure à la causalité ; cependant l'étude est de bonne qualité et au regard de la significativité cette hypothèse est raisonnable.

La réponse E est fausse : deux études sont rapportées à propos des seuls cancers du rein dont les auteurs rapportent qu'elles sont de faible qualité (taille des effectifs) : argument trop faible de causalité.

### Question 15

**Réponse exacte : A**

*Commentaires*

La réponse A est vraie : les auteurs voulaient étudier le lien entre polykystose rénale et survenue de cancer, or il n'est pas possible de randomiser le facteur « polykystose rénale ». Ce facteur ne peut s'étudier qu'au travers d'études observationnelles.

La réponse B est fausse : la cible est l'ensemble des patients présentant une polykystose rénale ; la population source et la population de résultat (ici confondue) sont limitée à la population de plus de 20 ans ; la polykystose dans sa forme autosomique récessive touche majoritairement une population mineure. Cette forme est plus rare mais présente des spécificités. Par ailleurs cette étude n'est conduite que parmi la population taïwanaise. Pour ces principales raisons il apparaît difficile d'extrapoler à l'ensemble des polykystoses.

La réponse C est fausse : le risque dans l'étude est réel ; il est significatif ; la causalité n'est pas démontrée, il n'y a pas lieu de proposer un dépistage spécifique hors du dépistage organisé en population sur le cancer du côlon à partir de 50 ans en France.

La réponse D est fausse : le risque est non significatif ; cependant les risques patientes présentant une polykystose rénale reste celui de la population générale pour laquelle un dépistage systématique est organisé au-delà d'un seuil d'âge ; il n'y a aucune raison de soustraire ces patientes au dépistage collectif.

La réponse E est fausse : l'étude a des qualités intrinsèques et permet d'émettre une hypothèse raisonnable ; mais c'est l'ensemble de toutes les études conduites sur le même sujet qui permettront d'évoquer un lien de causalité probable.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 4 Étude de cohorte (cohorte historique)

Aurélie Bourmaud, Mathieu Oriol

| | |
|---|---|
| **Titre de l'article** | **A cohort study on mental disorders, stage of cancer at diagnosis and subsequent survival** |
| **Auteurs** | Chang CK, Hayes RD, Broadbent MT, Hotopf M, Davies E, Møller H, Stewart R. A. |
| **Référence** | BMJ Open. 2014 ; 4(1):e004295 |
| **Lien vers l'article** | **https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3913023/** |

**1** **Les justifications énoncées par les auteurs pour avoir fait cette étude sont**

a. Certaines études ont démontré qu'il n'y avait pas d'association entre la mortalité liée au cancer et la présence d'une pathologie mentale sévère
b. De nombreuses études ont démontré une association entre la mortalité toutes causes et la présence d'une pathologie mentale sévère
c. De nombreuses études ont démontré une association entre la mortalité liée au cancer et la présence d'une pathologie mentale sévère
d. Des résultats contradictoires ont été retrouvés quant à l'association entre la mortalité liée au cancer et la présence d'une pathologie mentale sévère
e. Le niveau de preuve de ces études est insuffisant, il faut donc réaliser une étude expérimentale comparative pour conclure sur le lien de causalité entre la mortalité liée au cancer et la présence d'une pathologie mentale sévère

**2** **Les 3 phrases du paragraphe introduction, page 2, « First, to what extent... previously raised by Kisely et al » décrivent**

a. Des éléments de justification de l'étude
b. La population cible de l'étude
c. Les 2 stratégies testées dans l'étude
d. Les hypothèses de recherche des auteurs
e. Les objectifs de l'étude



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**3** **Quel est l'objectif principal de cette étude ?**

a. L'objectif de cette étude est d'identifier l'association entre les maladies mentales et le stade du cancer au moment du diagnostic ainsi que la survie, dans une population de patients diagnostiqués pour cancer
b. L'objectif de cette étude est d'identifier l'association entre les maladies mentales et le stade du cancer au moment du diagnostic ainsi que la survie, dans une population de patients diagnostiqués pour maladie mentale
c. L'objectif de cette étude est de comparer l'efficacité de la prise en charge de patients diagnostiqués pour maladie mentale sur le stade du cancer au moment du diagnostic ainsi que la survie, dans une population de patients diagnostiqués pour cancer
d. L'objectif de cette étude est de comparer l'efficacité de la prise en charge de patients cancéreux sur le stade du cancer au moment du diagnostic ainsi que la survie, dans une population de patients diagnostiqués pour maladie mentale
e. L'objectif de cette étude est de comparer l'efficacité de la prise en charge de patients cancéreux sur le stade du cancer au moment du diagnostic ainsi que la survie, dans une population représentative de la population générale régionale (London et autour)

**4** **Quel est le design de cette étude ?**

a. Cohorte prospective
b. Cohorte rétrospective
c. Observationnelle
d. Randomisée
e. Transversale

**5** **Il s'agit d'une étude observationnelle de cohorte rétrospective ou « historique ». Elle est ainsi intitulée parce que**

a. Il n'y a pas d'intervention du chercheur sur le choix des prises en charges et des expositions étudiées, selon un groupe où serait affecté le patient
b. Il n'y a pas de suivi des patients au cours du temps
c. Il y a appariement des groupes « maladie mentale » et « pas maladie mentale » sur les facteurs âge et genre
d. Il y a eu un suivi des patients au cours du temps, mais pour un autre objectif que celui de l'étude
e. L'étude a commencé après l'apparition des évènements étudiés : gradation du stade de cancer et décès lié au cancer/date de point

**6** **Les études observationnelles de cohorte rétrospectives ont habituellement comme caractéristiques**

a. D'avoir des biais d'information
b. D'avoir des perdus de vue
c. D'être peu coûteuses
d. De cumuler des faiblesses inhérentes aux études de cohorte prospectives et des faiblesses inhérentes aux études cas-témoin
e. De respecter les proportions vraies de malades/non malades



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**7** **Que vous évoque, page 3, paragraphe résultats, les phrases « Among them, 15 166 participants (34,7 %) had no information about their stage of cancer at diagnosis... After the exclusion of people without confirmed cancer stage information » ?**

a. Ces cas sont à l'origine d'un biais d'information
b. Ces cas sont à l'origine d'un biais de sélection
c. Ces cas sont à l'origine de données manquantes qu'il faudra prendre en compte dans l'analyse statistique
d. Le fait qu'il n'y ait pas de différence significative entre les caractéristiques de ces patients et de ceux n'ayant pas de données manquantes sur le stade permet de supposer que la population étudiée reste représentative de la population cible
e. Le fait qu'il y ait au total 34,7 % de patients n'ayant pas cette information permet de supposer que la population étudiée reste représentative de la population cible

**8** **Concernant le recueil des données relatif aux événements étudiés et aux expositions, dans les études observationnelles, des biais sont à rechercher**

a. Un biais d'attrition
b. Un biais d'information
c. Un biais de confusion
d. Un biais de rappel
e. Un biais lié à l'enquêteur

**9** **Concernant le recueil des données relatives au stade au diagnostic, issues du Thames Cancer Registry (une ou plusieurs réponses possibles)**

a. Il y a un biais de classement car les données proviennent de plusieurs établissements
b. Il y a un biais de classement car les données sont recueillies par plusieurs enquêteurs
c. Il y a un biais lié à l'enquêteur car les données recueillies sont celles des dossiers médicaux
d. Le biais de classement est partiellement contrôlé car la base de données finale est anonymisée
e. Le biais de classement est partiellement contrôlé car les enquêteurs qui remplissent la base sont formés

**10** **Concernant le recueil des autres données, appelées covariables, page 2, paragraphe « covariates included » on peut affirmer que**

a. Ce recueil est à l'origine d'un biais d'évaluation
b. Ce recueil est à l'origine d'un biais lié à l'enquêteur
c. Ces variables peuvent être des facteurs d'ajustement
d. Ces variables peuvent être des facteurs de confusion
e. Ces variables sont des données d'exposition



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**11** **Concernant les mesures d'association obtenues, quels sont les réponses justes (une ou plusieurs réponses possibles) ?**

a. La valeur des Hazard Ratios pour l'analyse de l'association « troubles bipolaires » et mortalité est en faveur du critère de Hill : force de l'association
b. La valeur des Hazard Ratios pour l'analyse de l'association « troubles bipolaires » et mortalité est en faveur du critère de Hill : fiabilité
c. Le fait qu'aucune relation linéaire n'ait été identifiée entre les 5 quintiles de déprivation et le stade au diagnostic n'est pas en faveur du critère de Hill : relation dose-effet
d. Le fait qu'aucune relation linéaire n'ait été identifiée entre les 5 quintiles de déprivation et le stade au diagnostic n'est pas en faveur du critère de Hill : spécificité de l'association
e. Le fait qu'il s'agisse d'une étude rétrospective n'est pas en faveur du critère de Hill : temporalité (la cause précède l'effet)

**12** **Concernant les résultats de l'étude (une ou plusieurs réponses possibles)**

a. Les résultats du tableau 1 correspondent à la description de la population de l'étude
b. Les résultats du tableau 1 correspondent à des résultats d'analyse univariée
c. Les résultats du tableau 2 correspondent à des résultats d'analyse univariée uniquement
d. Les résultats du tableau 2 correspondent à des résultats d'analyse multivariée uniquement
e. Les résultats du tableau 2 correspondent à des résultats d'analyse univariée et d'analyse multivariée

**13** **Concernant les résultats de l'étude (une ou plusieurs réponses possibles)**

a. L'ajustement sur certaines variables permet de contrôler les biais de confusion
b. La démence est un facteur de risque indépendant de stade avancé de cancer au diagnostic (OR = 1,47 $IC_{95\,\%}$ [1,01-2,14])
c. La schizophrénie est un facteur de risque indépendant de décès chez les patients avec un cancer (HR = 1,9 $IC_{95\,\%}$ [1,54-2,36])
d. Les résultats des OR et HR ajustés dans les tableaux 2 et 3 correspondent aux résultats des analyses multivariées
e. Tableau 3 : Les HR diminuent en valeur entre l'ajustement du modèle 1 et l'ajustement du modèle 2 car le facteur de confusion « stade au diagnostic » est pris en compte dans le modèle 2

**14** **Concernant les limites de l'étude (une ou plusieurs réponses possibles)**

a. Le fait que les données de maladie mentale soient mal renseignées avant 2006 est un inconvénient lié au caractère rétrospectif de la cohorte
b. Le fait que les données de stade au diagnostic de cancer soient mal renseignées à hauteur de 34,7 % est un inconvénient lié au caractère rétrospectif de la cohorte
c. Le fait qu'il n'a pas été possible d'ajuster sur les facteurs de style de vie est un inconvénient lié au caractère rétrospectif de la cohorte
d. Le fait que l'étude n'a pas réussi à identifier d'association significative entre les troubles de la personnalité et la mortalité est dû à un manque de puissance
e. Le fait que l'étude n'a pas réussi à identifier d'association significative entre les troubles de la personnalité et la mortalité est lié au caractère rétrospectif de la cohorte



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**15** **Concernant les conclusions de l'étude (une ou plusieurs réponses possibles)**

a. C'est parce que les Hazard Ratio restent significatifs entre le modèle 1 et le modèle 2 du tableau 3 que l'auteur dit « we found an association with worse survival after cancer diagnosis that was not explained by stage at presentation »
b. C'est parce qu'aucun résultat du tableau 2 n'est significatif que l'auteur dit « we found an association with worse survival after cancer diagnosis that was not explained by stage at presentation »
c. Réaliser une étude de cohorte prospective apporterait des données de plus haut niveau de preuve
d. Réaliser une étude de cohorte prospective apporterait des données sur le style de vie
e. Réaliser une étude de cohorte prospective apporterait plus de puissance statistique



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 1

**Réponses exactes : A – B – D**

*Commentaires* Les justifications d'une étude sont rapportées dans la partie introduction de l'article, de la première à l'avant dernière phrase (en général). Il s'agit d'une démonstration logique qui aboutit à une seule conclusion : l'hypothèse scientifique soulevée par les auteurs et traduite en objectif de recherche. Chaque phrase apporte un élément supplémentaire à la démonstration.

A. Cela est énoncé dans la phrase : « However, other studies found no association with cancer mortality for SMI as a whole or schizophrenia specifically ».

B. Cela est énoncé dans la phrase « Numerous studies have indicated a higher risk of all-cause mortality and shorter life expectancy for people with severe mental illness ».

C. Cela est contredit par la phrase : « Results from three population-based cohort studies showed significantly increased cancer mortality among people with schizophrenia for both genders, but some other studies reported that it occurred in men or women only. However, other studies found no association with cancer mortality for SMI as a whole or schizophrenia specifically and even a reduced risk was reported in one study ». L'ensemble des études faites sur le sujet ont des conclusions discordantes et non homogènes.

D. *Cf.* réponse à l'item C.

E. Ce n'est pas la conclusion des auteurs, qui ne proposent pas d'étude interventionnelle comparative, mais plutôt de faire une large étude de cohorte qui permettra de répondre à la question avec plus de puissance (plus de patients) : « Utilising a data linkage between a large secondary mental healthcare case register in southeast London and the regional cancer registry, we sought to investigate associations between mental disorder and disease stage at cancer diagnosis and subsequent survival. ». De plus, comme dans beaucoup d'études d'identification de facteurs de risque ou pronostiques, il n'est pas possible de faire un essai comparatif randomisé dans ce cas de figure.

## Question 2

**Réponses exactes : A – D**

**Inadmissible : C**

*Commentaires* Comme expliqué ci-dessus, le déroulement d'une introduction d'article est très structuré : après les éléments de la démonstration justifiant l'étude, l'auteur énonce les hypothèses de recherche qui sont la conclusion de cette démonstration. Puis après seulement il énonce la manière dont il va essayer de valider ces hypothèses : l'objectif de l'étude.

Attention tous les auteurs n'écrivent pas leurs hypothèses de recherche, elles sont parfois sous-entendues par celui-ci.

A. Les auteurs amènent ici un dernier argument issu de la littérature : « Then, an influence of mental disorders on cancer mortality in the absence of a clear effect on underlying risk could be explained by differences in treatment access, response and adherence, as previously raised by Kisely et al.

B. Non, celle-ci est la plupart du temps décrite dans l'objectif de l'étude.

C. Il ne s'agit pas d'une étude interventionnelle comparative, il n'y a donc pas de stratégies à comparer.

D. Les 2 hypothèses des auteurs sont :

- 1. la présence des troubles psychiatriques empêche la reconnaissance des signes cliniques précoces de cancer. Il y aurait donc une association entre maladie psychiatrique et stades



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

de cancer au diagnostic plus avancés pour ces patients : « First, to what extent might the reduced recognition of early cancer symptoms in people with mental disorders influence the stage of cancer at diagnosis ? » ;

- 2. il y a une association, indépendamment de la première, entre maladie psychiatrique et mortalité après diagnostic de cancer : « And, second, what is the role of mental disorders on mortality after cancer diagnosis if the issue of later presentation of cancer could be ruled out ? »

E. Non, ceux-ci sont notés dans la dernière phrase de l'introduction et suivent la règle classique de la composition d'un objectif de l'étude, qui n'est pas présente dans les 3 phrases citées dans la question 2.

**Question 3**

**Réponse exacte : A**

**Indispensable : A**

*Commentaires* L'objectif principal est enoncé dans la dernière phrase de l'introduction : « Utilising a data linkage between a large secondary mental healthcare case register in southeast London and the regional cancer registry, we sought to investigate associations between mental disorder and disease stage at cancer diagnosis and subsequent survival ». On y retrouve 3 des 4 éléments indispensables à la composition d'un objectif principal d'étude observationnelle :

- l'action du scientifique : identifier une association : « investigate associations » ;
- l'exposition étudiée : la maladie mentale : « mental disorder » ;
- les 2 événements étudiés : le stade du cancer au diagnostic : « disease stage at cancer diagnosis » et la survie subséquente : « subsequent survival » ;
- il manque le dernier élément qui est la population cible de l'étude, elle est sous-entendue et c'est au lecteur de l'identifier : il s'agit de l'ensemble des individus ayant eu un diagnostic de cancer. Si cette information n'est pas directement donnée dans l'objectif, elle est souvent présente dans l'introduction. Ici on en avait la certitude, ailleurs, en lisant la première phrase des résultats : « A total of 43 746 cancer cases were identified from TCR records » : C'est la population de l'étude.

Dans les études observationnelles, il est conseillé au lecteur de tracer le tableau de contingence de l'étude à partir des différents éléments de l'objectif et/ou de la partie méthode, afin de ne pas se perdre dans l'étude. Ici il y a 2 tableaux de contingence :

| **Tableau 1** | **Maladie +**<br>**Stade avancé du cancer au diagnostic** | **Maladie -**<br>**Stade localisé du cancer au diagnostic** |
|---|---|---|
| Exposition +<br>Maladie psychiatrique | | |
| Exposition –<br>Pas de maladie psychiatrique | | |
| **Tableau 2** | **Maladie +**<br>**Décès après le diagnostic de cancer** | **Maladie -**<br>**Pas de décès après le diagnostic de cancer** |
| Exposition +<br>Maladie psychiatrique | | |
| Exposition –<br>Pas de maladie psychiatrique | | |



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Les 3 réponses où il était question de comparer une efficacité sont complétement inadaptées ici : il s'agit d'une terminologie d'étude interventionnelle

### Question 4

**Réponses exactes : B – C**

**Inadmissible : D**

*Commentaires* Il s'agit d'une étude :

- observationnelle (par opposition à interventionnelle : pas d'intervention du scientifique qui orienterait le patient à adopter tel ou tel comportement comme prendre un médicament, suivre une stratégie thérapeutique, ou une action de santé) ;
- analytique : des associations entre des expositions et des maladies sont mesurées, au travers d'un suivi ;
- cohorte : un échantillon unique est constitué : patients ayant un diagnostic de cancer. Les fréquences d'événements et d'expositions ne sont pas choisies. À l'inverse d'une étude cas-témoin où les fréquences d'événements sont choisies (les cas « malades » et les témoins « non malades » sont choisis et appariés) ;
- rétrospective : l'étude commence après que le recueil de données sur les événements a été réalisé. Par opposition à prospective où l'étude commence avant que les événements n'aient encore eu lieu, donc n'aient été recueillis.

### Question 5

**Réponses exactes : A – D – E**

*Commentaires* La cohorte historique est une cohorte rétrospective. L'objectif de l'étude est déterminé *a posteriori*, après que le recueil de données portant sur les événements et les maladies n'ait eu lieu. Elle partage donc les biais des études rétrospectives « classiques ». Par contre les données utilisées ont bien été recueillies au cours d'un suivi (réalisé pour un tout autre objectif que celui de l'étude), ce qui permet de limiter partiellement les biais liés aux cohortes rétrospectives classiques.

### Question 6

**Réponses exactes : A – B – C – D – E**

*Commentaires* Les études observationnelles sont sources de biais importants qui sont à l'origine de leur moindre niveau de preuve. Il existe 3 types de biais associés aux études observationnelles : de sélection, d'information et de confusion.

A. Comme toutes les études observationnelles : le biais d'information est le biais qui se constitue au moment du recueil des données d'exposition et du recueil des données de maladie. C'est le fait de classer systématique dans la mauvaise case (E+ ou E-, M+ ou M-) un individu.

B. Comme toutes les études de cohortes où un suivi est pratiqué, plus le temps de suivi des patients est long, plus il y a de perdus de vue. Or les perdus de vue sont souvent plus malades et/ou plus exposés que les patients qui restent suivis. Ces perdus de vue sont à l'origine d'un biais de sélection : leur exclusion diminue la représentativité de l'échantillon final. Ils n'entrent pas dans le tableau de contingence.

C. Comme aucun suivi supplémentaire n'est prévu puisque l'étude se greffe sur une cohorte déjà existante mais pour un autre objectif, les coûts sont diminués d'autant, comme dans le cas d'une étude cas-témoin.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

D. Il y a le cumul des biais de perdus de vue (cohorte prospective) et l'absence du critère de Hill de temporalité : si les dates d'expositions ne sont pas recueillies, il est difficile d'établir la plausibilité temporelle (la cause précède l'effet) entre une exposition et la maladie (une maladie psychiatrique déclarée 15 ans avant un cancer vs une maladie psychiatrique déclarée 6 mois avant un cancer, par exemple, n'aura pas le même impact sur un cancer).

E. Comme toutes les études cohortes : la sélection de l'échantillon se fait sans choisir les proportions de M+ ou M- qui le constitue, seulement à partir d'une population remplissant les critères d'inclusion. Puis, on ne fait qu'observer la survenue de la maladie dans son histoire naturelle au sein de cette population. Les incidences naturelles de la maladie sont donc respectées.

## Question 7

**Réponses exactes : B – D**

*Commentaires* Les études observationnelles sont sources de biais importants qui sont à l'origine de leur moindre niveau de preuve. Il existe 3 types de biais dans ce type d'études : sélection – information – confusion.

Les biais de sélection font que la population étudiée (incluse dans l'étude) n'est pas représentative de la population cible. La population cible, c'est la population d'intérêt lorsque l'on pose les hypothèses d'une étude, c'est la population à laquelle on aimerait voir généraliser les résultats. À cause des mécanismes inhérents à la recherche (organisation, délais, refus), il est rare que la population incluse soit la même que la population cible. Il y a donc un biais de sélection.

Afin de contrôler ce biais de sélection (vérifier que la population incluse n'est pas trop différente de la population cible), il existe des méthodes de vérification : Il est recommandé de comparer – avec des tests statistiques – les caractéristiques qui ont été recueillies chez les patients qui ne sont pas (ou plus) dans l'étude aux caractéristiques de ceux qui sont (ou sont restés) dans l'étude.

- Si ces caractéristiques ne diffèrent pas significativement entre les 2 groupes, cela implique qu'*a priori*, sur ce qui a été recueilli, la population non incluse ressemble à la population incluse, et que donc celle-ci reste représentative de la population cible.
- Si les caractéristiques diffèrent, alors on dit qu'il y a un biais de sélection : les patients qui ne participent pas à l'étude sont exclus de l'étude de façon non aléatoire : leur exclusion est peut être liée à la maladie ou aux expositions testées dans l'étude. Dans ce cas il y a non-représentativité.

C. Faux car ces patients ne sont tout simplement pas étudiés dans l'étude.

E. Faux : on considère en général que moins de 10 % de patients perdus de vue ou sans données est un taux acceptable. Au-dessus de 10 % il y a un vrai risque de biais de sélection.

## Question 8

**Réponses exactes : B – D – E**

**Inadmissible : A**

*Commentaires* Nous avons vu que les études observationnelles sont à l'origine de 3 type de biais : sélection – information – confusion. Lorsque l'on s'intéresse aux biais liés au recueil des données, il s'agit des biais d'information (ou mesure).

Il existe de nombreux biais d'information, mais les plus courants sont les biais de rappel : le patient se rappelle plus ou moins bien (volontairement ou involontairement) de ses expositions ; les biais liés à l'enquêteur : l'enquêteur recueille différemment les données s'il est au courant du statut malade/non malade du patient, s'il est de bonne humeur un jour et pas le lendemain, et



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

s'il y a plusieurs enquêteurs, le recueil sera fait différemment selon leurs caractères ; les biais de classement : si des appareils de dosages biologiques sont étalonnés différemment, la valeur biologique mesurée sera différente d'un appareil à un autre.

A. Le biais d'attrition n'existe que dans les études interventionnelles.

C. Le biais de confusion n'est pas lié au recueil de données mais à la partie analyse statistique : c'est la présence de variables associées de façon détournée à la maladie étudiée.

### Question 9

**Réponses exactes : A – B – C – E**

*Commentaires* Les biais d'information (ou de mesure) sont à l'origine d'une estimation erronée de la mesure d'association entre une maladie et une exposition puisqu'ils répartissent les patients dans les 4 cases du tableau de contingence de façon erronée.

A. Un biais de classement est présent lorsque les données Exposé/Non Exposé ou Malade/Non Malade ne sont pas recueillies de la même façon pour tous les patients. Ainsi les pratiques des établissements vont impacter le recueil de la variable « stade au diagnostic ». Pour contrôler ce biais il faut centraliser et standardiser le mode de recueil.

B. De même un biais de classement est présent lorsque les données sont recueillies par différents individus, qui ne recueilleront pas de la même façon les variables. Le mode de contrôle du biais est le même : centralisation et standardisation.

C. En effet si les données sont recueillies dans les dossiers médicaux, c'est qu'un « enquêteur » les a colligées. Cet enquêteur est dans le cas des dossiers médicaux la plupart du temps un clinicien (interne ou médecin) : il y a donc un biais lié à l'enquêteur.

D. L'anonymisation des données n'est pas réalisée pour contrôler des biais mais pour des raisons éthiques.

E. *Cf.* B. la formation des enquêteurs permet la standardisation du recueil des données.

### Question 10

**Réponses exactes : B – C – D – E**

**Inadmissible : A**

*Commentaires* Ces variables sont des expositions. Leur association avec le stade du cancer et la survie sera testée. En tant qu'expositions, ces variables sont soit des potentiels facteurs de risque/ facteurs protecteurs, soit des facteurs de confusion. On ne le sait pas pour l'instant. Si ces variables sont considérées d'emblée par les auteurs comme des facteurs de confusion (parce que cela a déjà été démontré dans la littérature par exemple) alors les auteurs peuvent décider d'ajuster l'analyse multivariée sur ces variables. Les réponses C, D, E sont donc toutes justes.

A. Le biais d'évaluation n'existe que lorsque l'on fait une « évaluation » sur un « critère de jugement » : cela n'existe donc que dans les études interventionnelles.

B. À partir du moment où ces variables sont recueillies par un clinicien, celui-ci est considéré comme un enquêteur et donc comme à même de générer un biais d'information.

### Question 11

**Réponses exactes : C – E**

*Commentaires* Les critères de Hill permettent de juger de la plausibilité de la causalité de l'association entre un facteur et une maladie, association qui a été mise en évidence de façon statistique. Ils sont les suivant : force de l'association, relation dose-effet, la cause précède l'effet



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

(temporalité), spécificité de l'association (actuellement remise en cause), reproductibilité des résultats, plausibilité biologique, cohérence biologique, présence de données expérimentales. Il est indispensable d'étudier les résultats d'une étude observationnelle au travers du prisme des critères de Hill, ou une association fortuite pourrait être mal interprétée.

A. L'OR est trop proche de 1 pour être en faveur de ce critère.

B. Ce critère n'existe pas dans la liste des critères de Hill.

C. Vrai : la relation dose-effet implique qu'à valeur croissante de l'exposition corresponde une valeur croissante de la mesure d'association.

D. Ça n'a rien à voir. Qui plus est ce critère a été abandonné car invalide.

E. Vrai : c'est une des faiblesses des études rétrospectives : il n'y a pas la possibilité de vérifier que l'exposition précède bien la maladie, de façon cliniquement et physio-pathologiquement plausible. Ici il est juste vérifié que le diagnostic de maladie mentale précède le cancer, mais la juste temporalité ne l'est pas (6 mois avant un cancer ou 15 ans avant, le délai d'exposition à une maladie mentale n'aura pas le même impact sur le cancer).

## Question 12

**Réponses exactes : A – B – E**

*Commentaires* Le tableau 1 rapporte les caractéristiques des cas inclus dans l'étude (première colonne ou « all cases »), ainsi que celles des stades localisés de cancer et stades avancés de cancer (colonnes 2 et 3). Le tableau 1 rapporte aussi les résultats de l'analyse univariée, comparant les expositions citées en ligne (de « âge au diagnostic « à « score de déprivation ») au statut du cancer étudié : le stade au diagnostic. Les résultats des tests de cette analyse univariée sont donnés dans la colonne 4 ou « p-value ».

Le tableau 2 donne les résultats de tests d'association de l'événement « stade de cancer au diagnostic » aux variables d'expositions citées en lignes : les maladies mentales une par une. Il donne tout d'abord les résultats de l'analyse univariée dans la colonne « OR non-ajusté ». Puis les résultats de 2 modèles d'analyse multivariée sur les 2 colonnes suivantes « OR ajusté sur âge et genre » puis « OR tout ajusté ».

Donc A et B sont vrais pour le tableau 1 et E est vrai pour le tableau 2.

## Question 13

**Réponses exactes : A – C – D – E**

*Commentaires*

A. Il existe plusieurs façons de contrôler les biais de confusion dans les études observationnelles : soit au moment de la sélection de l'échantillon (appariement, stratification) ; soit au moment de l'analyse statistique : ajustement (on choisit les variables à mettre dans le modèle) ou analyse multivariée simple (c'est la significativité en univariée qui choisit les variables à mettre dans le modèle). Ici les auteurs ont choisi d'ajuster sur certaines variables.

B. L'OR = 1,47 est un OR brut, c'est-à-dire issu de l'analyse univariée. Une fois la variable analysée dans les 2 modèles multivariés, elle n'est plus significative : ce n'est pas un facteur de risque indépendant.

C. Vrai, il n'y a pas le 1 dans l'Intervalle de confiance à 95 %, c'est bien significatif en multivariée.

D. *Cf.* A.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

E. Afin d'éliminer l'effet du stade au diagnostic sur la mortalité, les auteurs ont ajusté leur modèle numéro 2 sur cette variable. Le peu de différence observée entre le modèle 1 et 2, confirme l'hypothèse comme quoi le stade au diagnostic n'influe que peu sur la mortalité dans cette population. L'explication de la surmortalité de cette population est ailleurs.

## Question 14

**Réponses exactes : A – B – C – D**

*Commentaires* Le caractère rétrospectif d'une cohorte (ou d'une étude cas témoin) est souvent associé à une qualité moindre du recueil des données que dans une cohorte prospective, construite pour répondre à l'hypothèse et à l'objectif de l'étude, pour 2 raisons :

- le recueil rétrospectif repose souvent sur des données clinico-biologiques recueillies dans les dossiers médicaux des patients. Plus ce recueil recule dans le temps et plus certaines données qui nous semblent indispensables actuellement manquent dans les dossiers, car non pertinentes cliniquement à l'époque ou pas encore existantes ;
- le fait que le cahier de recueil de données n'ait pas été construit pour l'objectif précis de l'étude, fait que certaines données indispensables dans ce contexte-là, manquent, car non indispensables pour l'objectif de la cohorte historique dans le contexte où elle a été créée.

A et B sont vraies et sont expliquées par la première raison. C'est vrai et est expliqué par la deuxième raison : les facteurs de style de vie n'étaient pas jugés pertinents dans le cadre des 2 cohortes historiques utilisées pour cette présente étude.

Par contre la non association entre mortalité et troubles de la personnalité semble être expliquée par un manque de puissance (8 et 8 patients concernés) et cela n'a rien à voir avec la qualité du recueil des données.

## Question 15

**Réponses exactes : A – C – D**

*Commentaires*

A. *Cf.* question 13 E.

B. Cela n'a rien à voir, le tableau 2 nous explique seulement qu'aucune maladie mentale n'expliquerait un stade plus avancé au diagnostic, même en tenant compte des caractéristiques du patient et de son statut socio-économique.

C. En effet construire son recueil de données en fonction de l'objectif de l'étude, avant de commencer l'étude, permet de recueillir les données pertinentes pour l'hypothèse testée, de bonne qualité et de façon systématique.

D. En effet si la littérature explique que le style de vie est une des expositions à étudier dans cette situation, ou si l'auteur s'y intéresse, le fait de réaliser une cohorte prospective lui permet de construire son recueil de données avec prise en compte de ces variables, avant de commencer l'étude.

E. Cela n'a rien à voir : les cohortes recueillent les chiffres observés d'exposition et de maladie. Que ce soit en rétrospectif ou en prospectif, l'auteur n'influencera pas les proportions observés. Si l'auteur veut augmenter la puissance statistique, il faut qu'il sélectionne une plus grosse population : mais cela est vrai en prospectif ou en rétrospectif.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 5 Étude cas-témoin

Adrien Guilloteau

| | |
|---|---|
| **Titre de l'article** | **Lifestyle, Family History, and Risk of Idiopathic Parkinson Disease: A Large Danish Case-Control Study** |
| **Auteurs** | Kenborg L, Lassen CF, Ritz B, Andersen KK, Christensen J, Schernhammer ES, Hansen J, Wermuth L, Rod NH, and Olsen JH |
| **Référence du n° de la revue** | Am J Epidemiol. 2015 ; 181(10):808–816 |
| **Lien vers le site** | https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4423523/ |

## 1 L'objectif de cette étude était d'évaluer

a. L'impact isolé de la consommation d'alcool sur le risque de survenue d'une maladie de Parkinson idiopathique (MPI)
b. L'impact isolé de la consommation de tabac sur le risque de survenue d'une MPI
c. L'impact isolé de la consommation da café sur le risque de survenue d'une MPI
d. L'impact isolé d'une histoire familiale de MPI sur le risque de survenue d'une MPI
e. L'existence d'une synergie entre la consommation de tabac et une histoire familiale de MPI sur le risque de survenue d'une MPI

## 2 Les cas

a. Ont été recrutés en population générale
b. Comprenaient 1 828 patients après applications des critères de sélection
c. Ont été initialement identifiés à partir de codes diagnostic de la Classification Internationale des Maladies, suite à une hospitalisation
d. Ont été définitivement inclus après une consultation qui a permis de confirmer le diagnostic de MPI
e. Devaient respecter un critère d'âge pour pouvoir être inclus dans l'étude

## 3 La population de témoins

a. A été recrutée en population générale
b. Devait être 5 fois plus nombreuse que celle des cas, à la fin du processus de sélection
c. Devait être appariée à la population des cas sur la consommation de café, d'alcool, de tabac et sur l'histoire familiale de MPI
d. Comportait 1 876 individus à la fin du processus de sélection
e. A été réduite en excluant les individus ayant été identifiés comme étant atteints de troubles neurologiques non reliés à la MPI, en s'appuyant sur les codes de la Classification Internationale des Maladies



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **Le recueil**

a. A été réalisé selon les mêmes modalités pour l'ensemble des personnes incluses
b. A été réalisé de 2008 à 2010
c. Des variables s'est fait en utilisant les mêmes unités de mesure entre cas et témoins
d. De la consommation de café s'est fait à l'aide d'une unité de mesure généralisable à l'ensemble de la population cible
e. De l'alcool a été fait en considérant l'ensemble des individus buvant moins d'un verre standard par semaine comme abstinent

**5** **Dans la partie analyse**

a. Le risque de première espèce a été fixé à 10 %
b. Il n'a pas été inclus de réflexion sur le nombre de sujet nécessaire
c. L'utilisation d'une régression logistique classique (non-conditionnelle) comme analyse principale a été justifiée par des difficultés d'appariement
d. Les auteurs déclarent intégrer les variables d'appariement dans l'analyse principale
e. Les auteurs intègrent l'effet de la consommation de cigarette à deux temps différents

**6** **La description de la population (table 1), permet de**

a. S'assurer que le processus de randomisation s'est bien déroulé, ce qui est le cas ici
b. S'assurer que le processus de randomisation s'est bien déroulé, ce qui n'est pas le cas ici
c. S'assurer que la sélection des individus a été correctement menée, ce qui est le cas ici
d. S'assurer que la sélection des individus a été correctement menée, ce qui n'est pas le cas ici
e. Déterminer à quelle population l'étude est généralisable

**7** **Le tableau 2 (« Odds Ratios for Idiopathic Parkinson Disease by Smoking Behavior, Denmark, 2008–2010 »)**

a. Montre une augmentation significative du risque de survenue d'une maladie de Parkinson chez les fumeurs par rapport aux abstinents, quelle que soit la méthode d'analyse
b. Montre une diminution significative du risque de survenue d'une maladie de Parkinson chez les fumeurs par rapport aux abstinents, quelle que soit la méthode d'analyse
c. Montre l'existence d'une relation dose-effet significative pour le nombre de cigarettes fumées par jour
d. Montre l'existence d'une relation dose-effet significative pour le nombre de paquets /année
e. Rappelle que l'analyse principale a été ajustée sur l'année de naissance et le sexe



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**8** **La figure 1 illustre**

a. Un risque de survenue de maladie de Parkinson environ divisé par deux chez les fumeurs ayant fumé plus de 30 ans
b. L'effet du nombre de paquets/année sur la survenue d'une maladie de Parkinson idiopathique
c. Le risque de survenue de la maladie de Parkinson idiopathique chez les non-fumeurs
d. L'interaction entre la consommation de cigarettes par jour et l'ancienneté du tabagisme sur le risque de Parkinson
e. Des relations effets-dose significatives de l'ancienneté du tabagisme au sein de chacune des catégories de nombre de cigarettes fumées par jours (soit « 0,1–9 », « 10–19 » et « ≥20 »)

**9** **En dehors de la consommation de tabac, cette étude a permis de montrer**

a. Une diminution du risque de survenue de MPI chez les consommateurs de café par rapport aux non consommateurs
b. L'existence d'une relation dose-effet de la consommation de café sur le risque de survenue de MPI
c. Une diminution du risque de survenue de MPI chez les consommateurs d'alcool par rapport aux abstinents
d. L'existence d'une relation dose-effet de la consommation d'alcool sur le risque de survenue de MPI
e. Une diminution du risque de survenue de MPI chez les patients ayant une histoire familiale de MPI

**10** **Le tableau 5 (« Odds Ratios for Idiopathic Parkinson Disease From Stratified and Joint Analyses of Smoking Status and Family History of Parkinson Disease, Denmark, 2008–2010 ») rapporte l'analyse stratifiée sur l'existence d'une histoire familiale de MPI puis sur le tabagisme (en deux catégories : « non-fumeur », « ex-fumeur ou fumeur actif ») et enfin l'analyse de l'effet joint de ces deux facteurs. Il permet de constater**

a. Un effet similaire de l'histoire familiale de MPI sur le risque de survenue d'une MPI quelle que soit la strate de tabagisme
b. Un effet différent du tabagisme sur le risque de survenue d'une MPI quelle que soit la strate d'histoire familiale de MPI
c. Une interaction entre l'effet du tabagisme et de l'histoire familiale de MPI sur la survenue d'une MPI
d. Une atténuation du risque de survenue de MPI liée à l'histoire familiale de MPI, chez les ex-fumeurs ou fumeurs actifs par rapport aux non-fumeurs
e. Que 20 modèles ont été nécessaires pour produire l'ensemble des odds-ratios du tableau



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**11** **Le phénomène de causalité inverse est abordé par les auteurs dans la discussion, il**

a. Représente l'existence d'une diminution du risque associée à un facteur explicatif (odd-ratio inférieur à 1)
b. Peut être la conséquence d'une modification de comportement des patients suite à un diagnostic de maladie
c. Améliore la crédibilité des résultats produits
d. Est limité par l'utilisation du recueil de l'exposition 10 ans avant les premiers symptômes
e. Est favorisé par l'utilisation du recueil de l'exposition 10 ans avant les premiers symptômes

**12** **Le biais de mémorisation**

a. Est essentiellement présent dans les études rétrospectives
b. Peut-être problématique dans le cadre d'études s'intéressant à des patients atteints de troubles neurologiques
c. Peut être limité par un recueil complémentaire dans l'entourage du malade
d. Est un biais différentiel jusqu'à preuve du contraire
e. Entraîne une surestimation des estimations produites (des odd-ratios dans cet article)

**13** **Le biais de classement dans les groupes**

a. A été en partie évalué par le retour à 50 dossiers patients
b. Est la conséquence de l'utilisation de critères diagnostiques clairement définis et validés dans la littérature scientifique
c. Correspond à l'existence d'un troisième facteur à la fois associé à l'évènement d'intérêt et au facteur d'exposition étudié et venant donc perturber la relation observée entre ces deux derniers
d. A entraîné le recrutement de patients plus graves que la population de patients atteints de MPI
e. Est, selon les auteurs, important dans cette étude

**14** **L'effet du tabac sur la survenue d'une MPI**

a. A une bonne validité externe
b. Est cohérent dans l'ensemble des analyses
c. Est causal, sur la base des résultats de cette étude
d. Peut être qualifié de protecteur
e. Est bien représenté par une relation effet-dose de la durée du tabagisme

**15** **Sur la qualité générale de l'article**

a. Le nombre de modèles réalisés contribue positivement à sa qualité
b. Du fait du nombre de patients recrutés, cette étude cas-témoins a le même niveau de preuve qu'une étude de cohorte
c. Les auteurs ont utilisé plusieurs méthodes pour limiter les différents types de biais
d. Le recrutement de la population de patients ne favorisait pas l'existence de biais de sélection
e. Les auteurs n'ont pas cherché à explorer une différence de participation entre les cas et les témoins



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 1

**Réponses exactes : A – B – C – D – E**

*Commentaires* L'objectif est classiquement mentionné à la fin de l'introduction : « In the present case-control study, we assessed the *individual and combined associations* of cigarette smoking, caffeine intake, and alcohol consumption with a family history of PD in a large group of patients with idiopathic Parkinson disease (IPD) ».

Pour cette étude Kenborg et al ont cherché à évaluer l'impact de 4 différents facteurs (consommation de tabac, d'alcool, de café, et histoire familiale de MPI), sur le risque de survenue d'une MPI, d'abord isolément, puis en synergie.

### Question 2

**Réponses exactes : C – E**

*Commentaires* La population des cas a été initialement identifiée sur la base d'un registre national hospitalier, répertoriant l'ensemble des séjours hospitaliers associés aux diagnostics posés lors de ces séjours. Les diagnostics sont codés sur la base de la Classification Internationale des Maladies (versions 8 et 10). De plus n'ont été retenus que les patients ayant été traités dans les 10 plus gros centres de traitement neurologiques. Le recrutement est ici exclusivement hospitalier. Après vérification sur dossiers médicaux (et non suite à une consultation) de la validité du diagnostic de maladie de Parkinson idiopathique et après applications des critères d'exclusion, la population finale de cas était de 1 808 patients.

### Question 3

**Réponses exactes : A – D – E**

**Inadmissible : C**

*Commentaires* La population de témoins a été recrutée en population générale sur la base d'un registre de population. Pour chaque cas, 5 témoins de même sexe et ayant la même date de naissance ont été tirés au sort dans ce registre, néanmoins au final un seul parmi ces 5 témoins devait être retenu.

### Question 4

**Réponses exactes : B – C**

*Commentaires* Pour la grande majorité de l'échantillon, le recueil a été fait par l'intermédiaire d'un entretien téléphonique, cependant pour les personnes ayant des difficultés d'audition, un recueil par courrier a été réalisé. Il n'est d'ailleurs pas rapporté la proportion respective de ces deux modes de recueil chez les cas et chez les témoins. Cette variable pourrait avoir de l'importance dans la mesure où la modalité de recueil peut influencer la réponse (ex : difficulté à avouer une forte consommation d'alcool en entretien par rapport à du déclaratif sur un questionnaire papier). L'évaluation de la consommation de café par tasse consommé par jour peut être problématique pour l'analyse et la généralisation des résultats. En effet au sein même de l'étude un individu pourra déclarer consommer une tasse, qui, en termes de caféine sera équivalente à deux tasses pour un autre individu. De plus il existe de grande variation au niveau international qui fait que la quantité moyenne de caféine par tasse varie considérablement d'un pays à l'autre. Ces points sont évoqués dans la discussion de l'article. Concernant la consommation d'alcool seuls les individus ayant consommé moins d'une unité d'alcool par semaine, pendant moins d'un an ont été considérés abstinents, et non l'ensemble des individus consommant moins d'une unité d'alcool par semaine.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 5

**Réponses exactes : B – C – D – E**

*Commentaires* Dans cet article, les analyses sont décrites de façon détaillée pour pouvoir expliciter les différentes analyses en regard des objectifs fixés. Il est mentionné que l'ensemble des analyses est réalisé avec « A 2-sided 5 % significance level », synonyme de risque alpha ou également de première espèce fixée à 5 %. La régression logistique conditionnelle est l'outil statistique permettant d'analyser des données d'étude cas-témoin avec appariement et ajustement. Il est clairement écrit au début du paragraphe sur la partie décrivant les analyses que l'analyse sera faite avec une régression logistique classique (non-conditionnelle, qui ne prend donc pas en compte le caractère apparié des données) du fait d'un déséquilibre entre les groupes cas et témoins. De ce fait les auteurs ont été obligés de réintégrer les variables d'appariement (année de naissance et sexe), dans l'analyse classique, car ceux-ci sont des facteurs de confusion majeurs et classiques. Enfin l'évaluation de l'effet de la consommation de cigarette se fait à deux temps : la consommation au moment du premier symptôme, et 10 avant le premier symptôme. Ce découpage se retrouve dans la partie résultats.

### Question 6

**Réponses exactes : C – E**

**Inadmissibles : A – B**

*Commentaires* La randomisation est le processus qui permet de répartir de façon aléatoire l'ensemble d'un échantillon dans deux groupes recevant chacun une stratégie thérapeutique différente. Elle ne s'applique qu'aux études interventionnelles (essentiellement des essais cliniques) et pas aux études observationnelles, telles les études cas-témoins. Ici les caractéristiques moyennes des deux groupes semblent très proches et le processus de sélection s'est donc probablement bien passé. Il faut néanmoins noter une petite disparité dans le lieu de résidence. La description de la population est également l'occasion de voir qu'elles sont ses caractéristiques de bases (sociodémographiques, cliniques...). Cela permet de déterminer à quel point les résultats de cette étude sont applicables à d'autres populations. A noter que s'il s'agit d'un élément très important pour une étude ayant un but descriptif, ça l'est un peu moins pour une étude à but analytique.

### Question 7

**Réponses exactes : B – D – E**

**Inadmissibles : A**

*Commentaires* Les résultats rapportés dans le tableau 2 montrent des odd-ratios systématiquement inférieur à 1, des intervalles de confiance ne comprenant jamais 1 (à deux exceptions près), avec l'utilisation comme référence des abstinents. Ce résultat semble donc plutôt robuste. On constate également des p de tendance linéaire significatif (inférieur au risque alpha fixé dans la partie méthodes), ce qui signifie l'existence d'une relation dose-effet pour le nombre d'année de consommation de cigarettes, le nombre de paquets-année et l'âge à l'arrêt de la consommation. Ce tableau rapporte plusieurs analyses, la principale, uniquement ajustée sur l'année de naissance et le sexe, et une analyse secondaire ajustée également sur l'âge aux premiers symptômes, le niveau d'éducation, la consommation de caféine, d'alcool, l'histoire familiale de maladie de Parkinson et l'urbanisation du lieu de résidence.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 8

**Réponses exactes : A – D – E**

*Commentaires* La figure 1 illustre la modification de l'effet de la consommation de cigarette par jour (correspondant aux trois catégories de l'axe des abscisses) en fonction de l'ancienneté de cette consommation (représentée par les 4 couleurs différentes). En mettant en relation ces deux variables, elle permet de constater s'il existe ou pas une différence d'évolution au sein des catégories. En l'occurrence elle ne paraît pas ici importante : en dehors des petites variations touchant la couleur blanche, les trois autres couleurs, donc catégories d'ancienneté du tabagisme suivent une évolution semblable dans les trois différentes catégories de nombre de cigarettes fumées par jour. En confirmation de ce point, la significativité des relations dose-effet (rapportée, comme pour le tableau 2 sous la forme de p de tendance linéaire) est semblable dans l'ensemble des catégories de nombre de cigarettes fumées par jour ($p = 0,04$ pour la catégorie 0,1-9, $p = 0,03$ pour la catégorie 10-19 et $p = 0,03$ pour la catégorie $\geq 20$).

## Question 9

**Réponses exactes : A – B – C**

*Commentaires* Le tableau 3 permet de voir l'effet de trois facteurs sur le risque de survenue de MPI, après ajustement : la consommation de café, la consommation d'alcool et l'existence d'une histoire familiale de MPI. La consommation d'alcool a été subdivisée en trois variables : consommation de bière, de vin et d'alcool forts. Ce tableau s'interprète de la même façon que la table 2. Elle rapporte notamment une diminution du risque de MPI chez les consommateurs de café, par rapport aux non-consommateurs (odd-ratio inférieur à 1, intervalle de confiance à 95 % ne comprenant pas 1) associé à l'existence d'une relation dose-effet (« p for trend » inférieur au risque de première espèce). Les auteurs affirment l'existence d'un effet plus conséquent de la bière par rapport aux deux autres boissons sur la base des données de cette étude. Cette affirmation est contestable dans la mesure où l'ensemble des intervalles de confiances associé à la consommation de bière (pour les deux types d'ajustement, avec ou sans le délai de 10 ans) comprennent la valeur de l'estimation liée à la consommation de vin. L'utilisation de catégories identiques (« abstinent », « 0,1-7 » « > 7 ») pour chaque type d'alcool est également contestable, car pour la catégorie de consommation la plus élevée les estimations sont très imprécises (les intervalles de confiance sont très larges) du fait de trop faibles effectifs, en particulier pour les alcools forts.

## Question 10

**Réponses exactes : A – D**

*Commentaires* Le tableau 5 répertorie l'ensemble des analyses permettant de mettre en relation les variables tabagisme et histoire familiale de MPI. Dans les analyses stratifiées on observe des odd-ratio similaires dans toutes les strates, pour les deux variables. Dans les analyses jointes on constate l'absence d'interaction (« P for interaction » supérieur au risque alpha), et donc une atténuation du risque de survenue du MPI liée à l'histoire familiale chez les fumeurs ou ex-fumeurs, par rapport aux non-fumeurs. Au total 14 modèles ont été nécessaires pour produire l'ensemble des odd-ratios de ce tableau : 1 pour chaque strate plus 1 pour la population globale, soit 12 pour les analyses stratifiées (6 pour chaque type d'ajustement) et 2 pour l'analyse jointe (1 pour chaque type d'ajustement). Il est important d'avoir une idée du nombre de modèles utilisés pour faire les analyses dans un article car cela a un impact sur la qualité des résultats. Le risque alpha représente le risque d'accepter à tort une hypothèse sur la base d'une différence uniquement due au hasard : à 5 %, il correspond à un risque de 1 sur 20. Plus



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

le nombre de modèles est grand plus le risque d'accepter à tort un résultat lié au hasard est élevé. Au total environ une cinquantaine de modèles différents sont rapportés dans cet article ce qui est à prendre en considération dans la lecture des résultats. Par ailleurs, il est important de regarder la significativité du facteur d'interaction globalement : s'il n'est pas significatif, il n'est pas nécessaire d'aller analyser finement les variations de l'odd-ratio pour chaque catégorie des deux variables.

### Question 11

**Réponses exactes : B – D**

*Commentaires* La causalité inverse est un biais particulièrement problématique dans l'analyse d'un facteur d'exposition sur la survenue d'une maladie. Il entraîne l'existence d'une relation significative entre facteur d'exposition et survenue d'une maladie, mais dont la nature causale est inverse, à savoir : c'est la maladie qui est à l'origine de la modification du facteur d'exposition et non le facteur d'exposition qui est à l'origine de la survenue de la maladie. La causalité inverse correspond donc le plus souvent à une modification du comportement du malade suite à l'apparition de la maladie, celle-ci pouvant se manifester à bas bruit par des changements d'habitudes de vie (habitudes alimentaires, sorties, activités physiques...). Dans l'exemple on peut parfaitement imaginer des modifications de la consommation de cigarettes, de café (par exemple gêné par une rigidité croissante), en amont du diagnostic de maladie de Parkinson, et avant même que le patient ne s'aperçoive de ces perturbations. Pour limiter son effet les auteurs ont choisi de recueillir les facteurs d'exposition dix ans avant les premiers symptômes avérés, ce qui permet de s'assurer que la maladie n'est pas à l'origine des comportements observés. Les auteurs de l'article rappellent également que certains facteurs causaux de la maladie peuvent également influencer des facteurs d'exposition, entraînant à nouveau une causalité inverse. Dans cette situation, il est très difficile de déterminer l'existence d'une causalité inverse, en particulier si ces facteurs sont génétiques. Ce biais existera donc principalement dans les études recherchant l'impact de maladie d'origines multifactorielles (principalement les maladies cardio-vasculaires et les cancers, mais aussi les maladies neurodégénératives) ayant un mode d'apparition lent et progressif. A noter, on parle souvent de relation inverse ou négative pour décrire les relations dans lesquelles un facteur explicatif diminue le risque de survenue d'une maladie. Mais il n'y a aucune relation avec le phénomène de causalité inverse.

### Question 12

**Réponses exactes : A – B – C – D – E**

*Commentaires* Le biais de mémorisation est essentiellement présent dans les études rétrospectives car il correspond à la différence entre la réalité passée et les données recueillies auprès du patient lors d'un entretien. Lors d'un recueil prospectif les données sont recueillies en même temps que la réalité se produit et elles ne sont donc pas soumises à ce biais. Le plus souvent ce biais est différentiel entre cas et témoins : les patients affectés et leurs entourages font souvent un effort de mémorisation plus important sur les causes potentielles de la maladie (et donc les facteurs d'exposition recueillies lors d'études) que des individus non affectés. Dans la mesure où le recueil fait appel à la mémoire, il est évident que les patients atteints de troubles neurologiques et particulièrement mémoriels vont générer des biais plus ou moins important dans le recueil des données. Pour ce faire il est toujours possible de se baser sur l'entourage, afin d'avoir une vision la plus complète et vérifiée possible du facteur d'exposition recueilli. Comme rappelé dans l'article le biais de mémorisation concernera le plus souvent des facteurs de risque plutôt que des facteurs protecteurs. Cependant il faut noter qu'il s'agit des facteurs de risque perçus par le patient et son entourage et non ceux avérés par la science. Le



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

tabagisme ayant une image plutôt négative sur la santé quelle que soit la maladie concernée, il est possible qu'il ait été ici envisagé comme responsable de la maladie par certains patients, en dépit des résultats de cette étude.

### Question 13

**Réponse exacte : A**

*Commentaires* Le biais de classement correspond à une famille de biais concernant à la fois la répartition des malades dans les groupes mais également l'évaluation des facteurs d'expositions. Comme pour le biais de mémorisation il peut être différentiel (différence de l'impact du biais entre les deux groupes). Les auteurs ont essayé de limiter ce biais en évaluant la réalité des codes diagnostiques à travers les dossiers médicaux, et en s'appuyant sur des grilles diagnostics validées. L'utilisation de ce type d'outils permet de s'assurer que les malades sont bien malades, et que les individus sains sont bien indemnes. Dans cette étude seul le diagnostic de la maladie a été confirmé ; l'absence de maladie (repose uniquement sur l'absence du code AVC). Le biais de classement au sein des groupes a néanmoins été partiellement évalué par l'intermédiaire d'une ré-évaluation indépendante des dossiers patients. Les auteurs ont jugé qu'avec 96 % de cohérences, le biais de classification était minime. L'existence d'un troisième facteur (ex : tabagisme) associé à la fois au phénomène étudié (ex : maladie cardio-vasculaire) et à un facteur d'exposition (ex : boire du café) qui viendrait perturber la relation entre ces deux derniers, correspond à un biais de confusion. Il y a effectivement eu un recrutement de patients atteints de MPI non représentatif de la population de malades, car ceux-ci ont initialement été identifiés en milieu hospitalier. Ce problème est variable selon les maladies, pour la maladie de Parkinson ce mode de recrutement ne permet pas de recruter les patients peu graves traités uniquement en ambulatoire. Néanmoins il s'agit d'un biais de sélection et non de classement.

### Question 14

**Réponses exactes : A – B – D – E**

*Commentaires* Dans l'introduction les auteurs dénombrent plusieurs études mettant en évidence un lien statistique entre tabac et maladie de Parkinson idiopathique, ce qui est en faveur d'une bonne validité externe ; l'item A peut donc être considérée ici comme juste. Cependant la qualité des études n'est pas discutée dans l'article, et il conviendrait de disposer d'une analyse critique de ces études pour pouvoir évaluer plus précisément la validité externe de cet effet. La réponse à ce type de question dépend donc de l'appréciation de la qualité de la discussion vis-à-vis de la littérature existante.

Sur les données de l'étude, l'ensemble des analyses convergent vers un effet protecteur (une relation inverse) du tabagisme sur la survenue d'une MPI. Cet effet semble également dépendant de la « dose » de tabagisme évalué par plusieurs facteurs (ancienneté du tabagisme, paquet-année, âge à l'arrêt du tabac pour les ex-fumeurs). Bien que l'existence d'une relation dose-effet fasse partie des critères de causalité de Hill, on ne peut pas affirmer de relation causale dans ce contexte. Les études cas-témoins sont soumises à trop de biais pour pouvoir affirmer un tel type de relation sur la base d'une seule étude. De ce fait elles auront plus souvent pour objectif d'identifier de potentiels facteurs de risque qui seront par la suite confirmés (ou non) par des études ayant un meilleur niveau de preuve : cohorte et méta-analyse principalement, rarement étude expérimentale.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 15

**Réponse exacte : C**

*Commentaires* Comme abordé dans la réponse à la question 10, le nombre de modèles affecte plutôt négativement la qualité d'une étude. Néanmoins dans le cadre d'études épidémiologiques à but exploratoire (cherchant à identifier de potentiels facteurs de risque, en opposition à confirmatoire, cherchant à démontrer l'existence et la taille d'un effet précisé) c'est une pratique acceptable. Et ce d'autant plus qu'une partie de ces modèles était indispensable pour analyser finement les interactions. Une étude cas-témoin, même bien menée ne peut avoir le même niveau de preuve qu'une cohorte du fait des nombreux biais inhérents à ce type de schéma (en particulier le biais de mémoire). Néanmoins les auteurs ont utilisé de nombreuses moyens pour essayer de limiter les biais de cette étude et faire en sorte que les résultats constatés soit les plus fiables possible, malgré le schéma utilisé. Ils ont également essayé d'évaluer l'ampleur des biais existants (ré-analyse des dossiers patients, comparaisons des témoins participants et non participants...). Ce dernier point permet d'évaluer dans quel sens le biais va orienter les résultats (s'il risque de les modifier), ce qui permet d'ajouter des résultats à la discussion.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 6 Étude cas-témoin

Nadia Oubaya, Bastien Boussat

| | |
|---|---|
| **Titre de l'article** | **Benzodiazepine use and risk of Alzheimer's disease: case-control study** |
| **Auteurs** | Billioti de Gage S, Moride Y, Ducruet T, Kurth T, Verdoux H, Tournier M, Pariente A, Bégaud B |
| **Référence** | BMJ. 2014 ; 349:g5205 |
| **Lien vers l'article** | http://www.bmj.com/content/349/bmj.g5205 |

**1** **Quel(s) étai(en) t le(s) objectif(s) de cette étude ?**

a. Étudier l'association entre les antécédents de consommation de benzodiazépines et la survenue de maladie d'Alzheimer
b. Étudier l'association entre l'anxiété et la survenue de maladie d'Alzheimer
c. Étudier l'association entre l'infarctus du myocarde et la survenue de maladie d'Alzheimer
d. Étudier l'association entre l'âge et la survenue de maladie d'Alzheimer
e. Étudier les facteurs pronostiques de maladie d'Alzheimer

**2** **Il s'agit d'une étude**

a. Randomisée
b. Expérimentale
c. Cohorte exposés/non exposée
d. Observationnelle
e. Transversale

**3** **Le choix d'une étude épidémiologique de type cas-témoin se justifie par**

a. Le long délai d'apparition de la maladie d'Alzheimer après l'exposition aux benzodiazépines
b. La rareté de l'exposition aux benzodiazépines dans la population générale
c. La possibilité d'étudier l'effet des benzodiazépines sur plusieurs pathologies
d. La possibilité d'étudier l'effet de plusieurs facteurs d'exposition
e. La possibilité d'effectuer un suivi longitudinal des cas

**4** **Quelle(s) est (sont) la (les) proposition(s) exacte(s) concernant les critères de sélection de la population de l'étude ?**

a. Pour être témoin, il ne fallait pas être atteint de maladie d'Alzheimer
b. Pour être témoin, il ne fallait pas avoir consommé de benzodiazépines
c. Pour être cas, il fallait avoir consommé des benzodiazépines
d. Pour être cas, il fallait être atteint de maladie d'Alzheimer
e. Les cas étaient issus de la base de données RAMQ



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**5** **Concernant les méthodes, quelle(s) est (sont) la (les) proposition(s) exacte(s) ?**

a. Les cas incidents ont été inclus dans l'étude, ce qui a permis de mieux contrôler la séquence temporelle
b. Le recrutement des cas permet de limiter un potentiel biais de survie sélective
c. Le recrutement des cas permet de qualifier l'étude de prospective
d. Le recrutement porte sur des témoins représentatifs des cas
e. Le recrutement porte sur des témoins comparables aux cas pour l'exposition étudiée (utilisation de benzodiazépines)

**6** **Chaque cas était apparié à 4 témoins. Par rapport à un appariement 1:1, cet appariement 1:4 permet de**

a. Mieux maîtriser les biais de confusion
b. Augmenter la puissance de l'étude
c. Augmenter la précision des résultats
d. Réduire les biais de sélection
e. Réduire l'étendue des intervalles de confiance

**7** **Le recueil de l'exposition aux benzodiazépines**

a. Concerne les traitements postérieurs à la date d'inclusion
b. Prend en compte les médicaments consommés dans les 20 ans précédant la date d'inclusion
c. Prend en compte les médicaments consommés dans les 10 ans précédant la date d'inclusion
d. Prend en compte les médicaments consommés entre 10 et 5 ans avant la date d'inclusion
e. Est limité sur une période définie afin de limiter le risque de non-respect la séquence temporelle

**8** **Les techniques utilisées dans cette étude visant à limiter les biais de confusion incluent**

a. L'appariement
b. L'ajustement
c. La stratification
d. La randomisation
e. La restriction de la population de l'étude

**9** **Les facteurs de confusion potentiels maîtrisés lors de l'analyse statistique incluent**

a. La consommation d'alcool
b. L'antécédent d'infarctus du myocarde
c. L'hypertension artérielle
d. L'âge
e. Le sexe



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**10** **Dans le tableau 2, concernant la comparaison entre les cas et témoins de la variable utilisation de benzodiazépines (utilisateur vs. non-utilisateur), quelle(s) méthode(s) les auteurs ont-ils pu utiliser ?**

a. Un test de Student
b. Un test du Chi 2
c. Un test du coefficient de corrélation
d. Une ANOVA
e. Un test du logrank

**11** **Concernant les résultats du tableau 2, quelle(s) est (sont) la (les) proposition(s) exacte(s) ?**

a. Les cas et les témoins ne différaient pas en termes d'antécédent de diabète grâce à l'appariement
b. Les cas et les témoins différaient en termes d'antécédent d'infarctus du myocarde ce qui montre que l'appariement n'a pas fonctionné
c. La proportion de patients ayant eu un AVC était significativement plus élevée chez les cas que chez les témoins
d. Il n'y avait pas d'intérêt à présenter les p pour les variables âge, sexe et durée de suivi car ils auraient forcément été inférieurs à 0,05
e. La proportion de patients ayant eu une hypercholestérolémie était significativement plus élevée chez les cas que chez les témoins

**12** **Concernant les méthodes et résultats, quelle(s) est (sont) la (les) proposition(s) exacte(s) ?**

a. Des rapports de cote bruts ont été présentés
b. Les auteurs ont pris en compte des éventuels facteurs de confusion
c. Les auteurs ont recherché une interaction entre dépression et maladie d'Alzheimer dans la consommation de benzodiazépines
d. Une analyse stratifiée a été présentée
e. Les auteurs n'ont pas retrouvé d'interaction entre la consommation de benzodiazépines et l'anxiété dans la survenue de maladie d'Alzheimer



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**13** **Concernant les résultats du 1er modèle multivarié du tableau 3, quelle(s) est (sont) la (les) proposition(s) exacte(s) ? On considérera que l'odds ratio est un bon estimateur du risque relatif.**

a. La cote d'exposition aux benzodiazépines chez les cas était plus élevée que la cote d'exposition aux benzodiazépines chez les témoins
b. Les sujets ayant été exposés à des benzodiazépines de longue demi-vie avaient un risque de survenue de maladie d'Alzheimer multiplié par 1,7 comparé aux sujets exposés à des benzodiazépines de courte demi-vie ou non exposés aux benzodiazépines
c. La cote de survenue de maladie d'Alzheimer chez les sujets ayant eu une densité d'exposition aux benzodiazépines > 180 PDD était supérieure à la cote de survenue de maladie d'Alzheimer chez les sujets ayant eu une densité d'exposition aux benzodiazépines nulle
d. L'analyse a retrouvé une augmentation relative du risque de survenue de maladie d'Alzheimer de 51 % chez les sujets exposés aux benzodiazépines comparé aux sujets non exposés aux benzodiazépines
e. L'exposition aux benzodiazépines était un facteur de risque de survenue de maladie d'Alzheimer, indépendamment de l'anxiété

**14** **Quel(s) est (sont) le(s) argument(s) de l'étude en faveur d'une relation de causalité entre exposition aux benzodiazépines et survenue de maladie d'Alzheimer ?**

a. L'existence d'une relation « dose-effet »
b. La cohérence avec les résultats d'études antérieures
c. L'analogie
d. La plausibilité biologique
e. La taille importante de l'échantillon

**15** **Considérant les résultats de cette étude, les recommandations à tirer en termes de santé publique incluent**

a. L'inscription d'un message d'avertissement sur les notices des benzodiazépines
b. Le retrait de l'autorisation de mise sur le marché pour les benzodiazépines
c. Une campagne à l'attention des médecins pour les alerter des dangers des benzodiazépines
d. Une restriction des indications des benzodiazépines
e. Il n'était pas possible de tirer des recommandations à partir de cette étude



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 1**

**Réponse exacte : A**

*Commentaires* L'objectif se trouve en fin d'introduction. Il s'agissait ici d'étudier l'association entre les antécédents de consommation de BZD (qui est la variable d'exposition) et la maladie d'Alzheimer (qui est le critère de jugement, ou statut cas-témoin). L'anxiété et l'infarctus du myocarde constituent des variables d'ajustement, c'est-à-dire qu'il s'agit de potentiels facteurs de confusion qui sont pris en compte dans l'analyse, et on ne cherche pas à étudier l'association entre ces variables et la maladie d'Alzheimer. L'âge est un facteur d'appariement, on ne cherche donc aucunement à étudier l'association entre l'âge et la maladie d'Alzheimer. En effet, les facteurs sur lesquels cas et témoins sont appariés sont, de fait, rendus identiques entre cas et témoins, il n'est donc pas possible d'étudier ensuite l'association entre les facteurs d'appariement et le statut cas-témoin.

Les facteurs pronostiques de la maladie d'Alzheimer seraient les facteurs d'événements délétères de santé (tels que l'institutionnalisation ou le décès) survenant chez les patients atteints de maladie d'Alzheimer alors que les facteurs prédictifs de maladie d'Alzheimer sont les facteurs de risque de survenue de maladie d'Alzheimer.

**Question 2**

**Réponse exacte : D**

**Indispensable : D**

**Inadmissible : A**

*Commentaires* Il s'agissait d'une étude cas-témoin, dont l'objectif était d'étudier la relation entre l'utilisation de benzodiazépine (exposition) et la maladie d'Alzheimer. Les études cas-témoins sont des enquêtes épidémiologiques, observationnelles, comparatives. Pour différencier une étude cas-témoin d'une étude de cohorte, on peut identifier si les sujets sont recrutés selon le statut malade/non malade (cas-témoin) ou exposé/non-exposé (cohorte). A- attention à ne pas confondre sélection aléatoire dans une population (random sample en anglais) et étude randomisée (l'allocation à un traitement ou une intervention est tirée au sort).

**Question 3**

**Réponses exactes : A – D**

*Commentaires* Les études cas-témoins sont adaptées à l'étude de maladies rares pour lesquelles le délai depuis le début de l'exposition est relativement long, situation dans laquelle une enquête de cohorte nécessiterait une population de grande taille suivie sur plusieurs années. Elles permettent également de prendre en compte plusieurs facteurs d'exposition simultanément. Les études cas-témoins ne permettent pas d'effectuer un suivi longitudinal, et le recueil des expositions est toujours rétrospectif.

B. Outre le fait que l'exposition aux benzodiazépines dans la population générale ne peut être qualifiée de « rare », l'étude de la relation entre une exposition rare et une maladie fera préférer le choix d'une cohorte.

C. Dans les études cas-témoins, les sujets sont recrutés selon la présence ou non d'une pathologie. On peut donc étudier l'effet de plusieurs facteurs d'exposition sur cette pathologie, mais il n'est pas possible d'étudier leurs effets sur d'autres pathologies.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

■ **Question 4**

**Réponses exactes : A – D – E**

*Commentaires* Les cas étaient les sujets atteints de maladie d'Alzheimer et les témoins étaient les sujets indemnes de maladie d'Alzheimer. Le statut-cas témoin correspond au critère de jugement (= variable à expliquer = variable dépendante = événement = outcome ou endpoint en anglais). La consommation de benzodiazépines (BZD) est la variable d'exposition (= variable explicative = variable indépendante). Et l'étude cas-témoins consiste à comparer entre les cas et les témoins la proportion de sujets ayant eu l'exposition, donc le statut cas et témoin est défini indépendamment de l'exposition.

■ **Question 5 :**

**Réponses exactes : A – C**

*Commentaires* Seuls les cas incidents ont été considérés (pour être cas, il fallait avoir un premier diagnostic de maladie d'Alzheimer). La date index correspond pour les cas à la date de l'événement et pour les témoins à la date de l'événement du cas auquel ils sont appariés. Le recrutement portait sur des cas incidents. En effet, les cas étaient inclus si un premier diagnostic de maladied'Alzheimer était enregistré durant la période du plan entre 2001 et 2009. L'avantage de recruter des cas incidents est de limiter un biais de sélection par la survie sélective. Les patients sont inclus quelle que soit leur durée de survie, et on limite la sélection pouvant être liée au facteur de risque étudié, dans le cas où le rôle de celui-ci est, en plus de provoquer la maladie, de diminuer la durée de survie. Les témoins ne sont par définition pas comparables aux cas. A minima, ils ne sont pas comparables de par l'absence de la maladie étudiée.

■ **Question 6**

**Réponses exactes : B – C – E**

*Commentaires* Sélectionner un groupe témoin supérieur au groupe de cas conduit à une augmentation de la puissance de l'enquête. La puissance est similaire au pouvoir grossissant d'un microscope. Elle est fonction du nombre de sujets, de la taille de la différence réelle (ou force de l'association réelle), et du risque de première espèce.

A. Apparier 4 témoins pour 1 cas ne permettra pas une meilleure maîtrise des biais de confusion. L'appariement permet de rendre les cas comparables aux témoins sur les facteurs d'appariement (facteurs de confusion potentiels ciblés lors de la planification de l'enquête), quel que soit le nombre de témoins appariés.

C-E. La largeur de l'intervalle de confiance reflète la précision de l'estimation d'un paramètre statistique sur un échantillon. Plus le nombre de sujets est grand, plus l'intervalle de confiance est étroit, et plus les résultats sont précis.

■ **Question 7**

**Réponses exactes : D – E**

*Commentaires* Dans une enquête cas-témoin, le recueil des expositions est toujours rétrospectif. Dans cette enquête, les auteurs ont recueilli l'exposition aux benzodiazépines dans la fenêtre comprise entre 10 et 5 ans avant la date d'inclusion. Cette décision a pour but de limiter une des faiblesses inhérentes aux études cas-témoin : le non-respect de la séquence temporelle, la séquence chronologique étant difficile à établir.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 8

**Réponses exactes : A – B – E**

*Commentaires* Plusieurs techniques sont possibles pour limiter les biais de confusion. Au moment de la planification de l'étude, il est possible de randomiser les groupes, restreindre la population de l'étude, équilibrer la distribution de facteurs de confusion entre les groupes en appariant. La randomisation n'est cependant pas réalisable lorsque l'on se trouve en situation d'observation, comme c'est le cas d'une étude cas-témoin. Au moment de l'analyse statistique, il est possible de maîtriser des facteurs de confusion non pris en compte lors de la constitution de l'échantillon, en modélisant la relation entre la maladie et l'exposition, indépendamment de l'exposition à des facteurs de confusion.

A. Appariement sur le sexe, l'âge et la durée de suivi antérieur à la date d'inclusion.

B. Régression logistique multivariable (model 1 et model 2), étudiant l'association entre l'utilisation de la benzodiazépine et la maladie d'Alzheimer indépendamment de la présence de cofacteurs.

E. En réalisant leur étude sur les personnes âgées de plus de 65 ans, les auteurs ont donc restreint leur étude sur une catégorie d'âge précise. L'âge jeune peut en effet être considéré dans la maladie d'Alzheimer comme une catégorie rare du facteur d'exposition « âge ». De plus, les cas de maladie d'Alzheimer chez les jeunes peuvent être considérés comme des formes particulières de cette maladie.

## Question 9

**Réponses exactes : B – C**

*Commentaires* Il fallait ici repérer les facteurs introduits dans les modèles de régression logistique multivariable, technique qui permet de prendre en compte les facteurs de confusion potentiels au moment de l'analyse statistique.

A. La consommation d'alcool n'était pas recueillie dans cette enquête.

B – C. Ces deux facteurs étaient introduits dans le modèle de régression logistique.

D – E. L'âge et le sexe étaient maîtrisés lors de la planification de l'enquête, et non lors de l'analyse statistique. On remarque bien la distribution comparable de ces 2 facteurs dans le tableau 2. Le terme « Matched » signifie que ces 2 facteurs étaient des facteurs d'appariement, il ne convenait donc pas de les introduire dans les modèles.

## Question 10

**Réponse exacte : B**

*Commentaires* La comparaison entre les cas et témoins de la variable utilisation de benzodiazépines dans le tableau 2 est une analyse univariée de l'association entre une variable qualitative binaire (cas vs. témoins) et une autre variable qualitative binaire (utilisateur de BZD vs. non-utilisateur de BZD). Le test à utiliser est donc le test du Chi 2 de Pearson (utilisable également si les variables sont qualitatives à plus de 2 classes).

Le test de Student étudie l'association entre une variable qualitative binaire et une variable quantitative. L'ANOVA (analyse de variance) étudie l'association entre une variable qualitative à plus de 2 classes et une variable quantitative. Le test du coefficient de corrélation (Pearson en paramétrique et Spearman en non paramétrique) étudie l'association entre deux variables quantitatives. Le test du logrank étudie l'association entre une variable qualitative (binaire ou à plus de 2 classes) et une variable qualitative binaire avec prise en compte du délai de survenue.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 11

**Réponse exacte : E**

*Commentaires* L'appariement rend les cas et témoins comparables concernant les critères d'appariement (ici âge, sexe, et durée de suivi) donc le p du test de comparaison d'un critère d'appariement entre cas et témoins sera non significatif (≥ 0,05) mais il peut y avoir des différences entre cas et témoins sur les autres critères (différence significative si p < 0,05 et on regarde les proportions retrouvées dans notre population d'étude pour connaître le sens de la différence).

## Question 12

**Réponses exactes : A – B – E**

*Commentaires* Le tableau 3 présente des rapports de cote (=odds ratios) bruts issus de l'analyse univariée et des rapports de cote ajustés issus de l'analyse multivariée. Les auteurs ont recherché des interactions entre la consommation de BZD et respectivement anxiété, dépression et insomnie, dans la survenue de maladie d'Alzheimer. Ils n'ont retrouvé aucune interaction significative (interaction = interaction = effect modification en anglais). Tous les tests d'interaction étaient non significatifs (p≥0,05). Les phénomènes d'interaction peuvent être étudiés en faisant une analyse stratifiée (en stratifiant sur le facteur d'interaction). Par exemple si on souhaite étudier l'interaction entre anxiété et consommation de BZD sur la survenue de maladie d'Alzheimer, on stratifie sur la variable anxiété et on étudie l'association entre consommation de BZD et survenue de maladie d'Alzheimer dans la strate (sous-groupe) présence d'anxiété et dans la strate absence d'anxiété. Les auteurs n'ont pas présenté d'analyse stratifiée dans cette étude.

## Question 13

**Réponses exactes : A – C – D**

*Commentaires* L'odds ratio est le rapport de cote. Il s'agit du rapport de la cote d'exposition chez les malades sur la cote d'exposition chez les non malades (ce qui équivaut au rapport de la cote de maladie chez les exposés sur la cote de maladie chez les non exposés). La cote est un rapport de 2 probabilités complémentaires (dont la somme fait 1).

$$OR = \frac{\text{Cote d'expostion chez } M+}{\text{Cote d'expostion chez } M-} = \frac{\dfrac{P(E+/M+)}{P(E-/M+)}}{\dfrac{P(E+/M-)}{P(E-/M-)}} = \frac{\dfrac{a/(a+c)}{c/(a+c)}}{\dfrac{b/(b+d)}{d/(b+d)}} = \frac{\dfrac{a}{c}}{\dfrac{b}{d}} = \frac{ad}{bc}$$

$$OR = \frac{\text{Cote d'expostion chez } E+}{\text{Cote d'expostion chez } E-} = \frac{\dfrac{P(M+/E+)}{P(M-/E+)}}{\dfrac{P(M+/E-)}{P(M-/E-)}} = \frac{\dfrac{a/(a+b)}{b/(a+b)}}{\dfrac{c/(c+d)}{d/(c+d)}} = \frac{\dfrac{a}{b}}{\dfrac{c}{d}} = \frac{ad}{bc}$$

| | M+ | M- |
|---|---|---|
| E+ | a | b |
| E- | c | d |

Pour mémoire, le risque relatif est : $RR = \dfrac{P(M+/E+)}{P(M+/E-)} = \dfrac{a/(a+b)}{c/(c+d)}$ ;

Il n'est pas calculable dans une étude cas-témoins car la proportion de cas et témoins est fixée par l'étude.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

L'item B est faux car le risque de survenue de maladie d'Alzheimer est multiplié par 1,7 comparé aux sujets non exposés aux BZD. La mesure d'association vaut toujours par rapport à la classe de référence et non par rapport à toutes les autres classes.

La différence relative de risque = RR-1, elle peut être positive (augmentation relative du risque) ou négative (réduction relative du risque).

La légende du tableau précise les facteurs sur lesquels un ajustement a été fait dans le modèle 1 et dans le modèle 2. Le modèle 1 n'était pas ajusté sur l'anxiété.

## Question 14

**Réponses exactes : A – B – D**

*Commentaires* Il s'agit des critères de Bradford Hill (critères en faveur d'une relation de causalité dans une étude observationnelle). Les critères retrouvés ici sont : relation dose effet (plus la dose est élevée plus la force de l'association est grande), cohérence avec les résultats d'études antérieures et plausibilité biologique. La taille de l'échantillon ne fait pas partie des critères de Bradford Hill. L'analogie est un des critères mais il n'est pas retrouvé dans cette étude. Analogie : il existe une association entre une exposition analogue et l'événement. Ex : association entre benfluorex et valvulopathies cardiaques ; il existait une association entre des coupe-faim analogues au benfluorex et la survenue de valvulopathies.

## Question 15

**Réponse exacte : E**

*Commentaires* Cette étude cas-témoin, retrouvait une relation significative entre l'utilisation de benzodiazépine et la maladie d'Alzheimer. Si cette étude a de nombreux points forts (large échantillon représentatif de 98 % de la population générale > 65 ans du Québec, recrutement de cas incidents, étude de la relation dose-effet...), son niveau de preuve reste faible compte tenu de son design cas-témoin, et elle comporte des points faibles importants (biais protopathique non écarté, pas d'évaluation directe des patients, séquence temporelle difficile à établir avec certitude). Avant d'émettre des recommandations en termes de santé publique, des enquêtes supplémentaires sont nécessaires afin d'établir la preuve scientifique de ce lien entre benzodiazépine et maladie d'Alzheimer.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 7 Étude cas-témoin (nichée dans une cohorte)

David Hajage, Sarah Feldman, Lisa Belin

| | |
|---|---|
| **Titre de l'article** | **Case–Control Study of Human Papillomavirus and Oropharyngeal Cancer** |
| **Auteurs** | D'Souza G, Aimee RK, Raphael V, Pawlita M, Fakhry C, Koch WM, Westra WH, Gillison ML |
| **Référence** | New Engl J Med 2007 ; 356 (19):1944-56 |
| **Lien vers l'article** | http://www.nejm.org/doi/full/10.1056/NEJMoa065497 |

**1** **L'objectif principal de cette étude était**

a. De déterminer l'existence d'une association entre l'infection à HPV et le cancer épidermoïde de l'oropharynx
b. De déterminer les facteurs prédictifs du cancer épidermoïde de l'oropharynx
c. De déterminer l'existence d'une association entre les pratiques sexuelles « à risque » et le cancer épidermoïde de l'oropharynx
d. De déterminer l'existence d'une association entre le statut alcolo-tabagique et le cancer épidermoïde de l'oropharynx
e. De déterminer l'intérêt d'une campagne de vaccination contre l'HPV sur l'incidence du cancer épidermoïde de l'oropharinx

**2** **Les caractéristiques définissant le mieux cette étude sont**

a. Analytique
b. Expérimentale
c. Cas-témoins
d. Quasi-expérimentale
e. Randomisée

**3** **Dans cette étude, les sujets atteints d'un cancer de l'oropharynx (les « cas ») étaient**

a. Des cas prévalents
b. Des cas incidents
c. Des cas inclus juste après le diagnostic de leur maladie
d. Des cas déjà diagnostiqués depuis plus ou moins longtemps au moment de leur inclusion
e. Toutes les réponses précédentes sont fausses



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **Par rapport à une étude réalisée sur des cas prévalents, l'(les) avantage(s) d'un recrutement de cas incidents sont**

a. Une durée d'inclusion plus courte
b. Une chronologie exposition – maladie moins litigieuse
c. Une clause d'ignorance plus facile à mettre en œuvre
d. Un biais de sélection moins important
e. Des exclusions secondaires post-randomisation moins nombreuses

**5** **Les caractéristiques des cas et des témoins en termes d'âge et de sexe (présentées dans le tableau 1) sont identiques**

a. C'est attendu du fait de l'ajustement effectué (analyse multivariée)
b. C'est attendu du fait de l'appariement effectué
c. C'est attendu du fait de la stratification effectuée
d. C'est attendu du fait de la restriction effectuée
e. Toutes les propositions précédentes sont fausses

**6** **Les cas et les témoins sont appariés sur l'âge (± 5 ans) et le sexe. Cet appariement induit**

a. Une diminution du biais de mémoire
b. Une diminution du biais d'avance au diagnostic
c. Une diminution du biais de confusion
d. Une diminution du biais protopathique
e. Une diminution du biais d'indication

**7** **Les auteurs n'ont pas apparié les cas et les témoins sur le statut alcoolique. Concernant l'évaluation de l'association entre infection à HPV et cancer de l'oropharynx, apparier sur le statut alcoolique**

a. N'aurait eu aucun impact concernant l'évaluation de cette association
b. Aurait possiblement facilité l'évaluation de cette association
c. Aurait possiblement compliqué l'évaluation de cette association
d. Aurait permis de prendre en compte ce facteur de confusion potentiel
e. N'aurait pas permis de prendre en compte ce facteur de confusion potentiel

**8** **Les auteurs n'ont pas apparié les cas et les témoins sur le fait d'avoir ou non un comportement sexuel « à risque ». Concernant l'évaluation de l'association entre infection à HPV et cancer de l'oropharynx, apparier sur les comportements sexuels « à risque »**

a. N'aurait eu aucun impact concernant l'évaluation de cette association
b. Aurait possiblement facilité l'évaluation de cette association
c. Aurait possiblement compliqué l'évaluation de cette association
d. Aurait permis de prendre en compte ce facteur de confusion potentiel
e. N'aurait pas permis de prendre en compte ce facteur de confusion potentiel



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**9** **Les cas atteints d'un cancer de l'oropharynx mais ayant refusé de participer à cette étude étaient plus souvent des femmes (p = 0,001). Cela induit**

a. Une diminution de l'efficacité de l'appariement
b. Une diminution du risque d'être atteint d'un cancer de l'oropharynx
c. Une diminution de la validité externe de l'étude
d. Une diminution de la représentativité de l'échantillon
e. Toutes les réponses précédentes sont fausses

**10** **Les éléments du schéma de l'étude présents dans l'article diminuant le risque de biais de mémoire différentiel entre les cas et les témoins sont**

a. La nature de l'exposition principale d'intérêt (infection à HPV mesurée en laboratoire)
b. Le recueil de l'exposition en aveugle de l'hypothèse de l'étude
c. Le recueil de l'exposition en aveugle du statut cas ou témoin
d. L'utilisation d'analyses multivariées
e. L'utilisation d'auto questionnaires

**11** **Les mesures d'association entre exposition et maladie rapportées dans l'article sont des odds ratios. Les autres mesures d'association calculables dans cette étude sont**

a. Le risque relatif
b. Le hazard ratio
c. La différence des risques
d. La différence de moyennes
e. Toutes les réponses précédentes sont fausses

**12** **Le tableau 3 rapporte les résultats principaux de l'étude concernant l'association entre infection à HPV et cancer de l'oropharynx. L'exposition est mesurée de différente façon (infection sérologique, infection orale, etc.), mais l'odds ratio associé aux sujets HPV négatif est toujours égale à 1. Cela signifie**

a. Que les sujets sans infection à HPV n'ont aucun risque de cancer de l'oropharynx
b. Qu'aucun cancer de l'oropharynx n'a été observé chez des sujets sans infection à HPV
c. Que la puissance de l'étude n'est suffisante que dans le sous-groupe des sujets infectés à HPV
d. Qu'il s'agit de la classe de référence
e. Qu'il existe probablement des facteurs de confusion non pris en compte

**13** **Toujours dans le tableau 3, le degré de signification associé au statut sérologique à HPV-16 L1 n'est pas rapporté par les auteurs**

a. Car il est supérieur à 0,05
b. Car il n'est pas nécessaire pour conclure sur la significativité de l'association
c. Car il ne peut être calculé du fait des facteurs de confusion
d. Car il ne peut être calculé dans une étude cas-témoins
e. Car la puissance de l'étude était insuffisante



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**14** **Quelles sont les informations apportées par le tableau 3 concernant la prévalence du cancer de l'oropharynx en population générale ?**

a. La prévalence du cancer de l'oropharynx est d'environ 50 %
b. La prévalence du cancer de l'oropharynx est d'environ 33 %
c. Le cancer de l'oropharynx est un cancer fréquent
d. Le cancer de l'oropharynx est un cancer rare
e. Toutes les réponses précédentes sont fausses

**15** **Concernant la vaccination anti-HPV systématique des filles et des garçons, les résultats de cette étude**

a. Permettent de la recommander car le niveau de preuve est élevé
b. Permettent de la recommander car les associations (OR) rapportées sont très élevées
c. Permettent de la recommander car les facteurs de confusion connus ont été pris en compte
d. Permettent de la recommander car l'échantillon était représentatif
e. Toutes les réponses précédentes sont fausses



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

■ **Question 1**

**Réponse exacte : A**

*Commentaires* L'objectif principal de l'étude est donné à la fin de l'introduction : « provide additional support for a causal association between HPV and oropharyngeal cancers ». Le titre de l'article est également très explicite. Attention de ne pas confondre l'objectif principal et la perspective de l'étude concernant l'utilité d'éventuels programmes de vaccination anti HPV (« and might guide future cancer-prevention program... ») qui pourraient être envisagés à l'issue de l'étude.

■ **Question 2**

**Réponses exactes : A – C**

**Indispensable : C**

**Inadmissible : E**

*Commentaires* Il s'agit d'une étude cas-témoins (réponse C vraie). Il s'agit d'une étude observationnelle, c'est-à-dire d'une étude où l'exposition d'intérêt est observée et non attribuée aux sujets selon un schéma expérimental. Il ne s'agissait donc pas d'une étude expérimentale (réponse B fausse) randomisée (réponse E fausse) ou quasi-expérimentale (les études avant-après ou ici-ailleurs) (réponse D fausse). Une étude observationnelle peut avoir une vocation descriptive (étude d'incidence ou de prévalence d'une maladie par exemple) ou analytique (synonyme : étiologique, étude cherchant à évaluer l'existence d'une association entre une exposition et une maladie) (réponse C). Attention, il s'agit bien d'une étude cas-témoins, même si celle-ci est nichée dans une étude de cohorte (première ligne de la section « Methods » : « nested within a longitudinal cohort study ».

■ **Question 3**

**Réponses exactes : B – C**

*Commentaires* Les cas sont issus d'une cohorte longitudinale monocentrique (à l'hôpital Johns Hopkins de Baltimore) de sujets nouvellement diagnostiqués d'un cancer de l'oropharynx (premier paragraphe de la section « Methods »). Les cas sont donc inclus au moment où ils deviennent malades, ce qui constitue la définition d'une étude cas-témoins réalisée sur cas incidents (réponses B et C vraies).

■ **Question 4**

**Réponses exactes : B – D**

*Commentaires* Le premier avantage d'une étude cas-témoins sur cas incidents (nouvellement diagnostiqués, inclus au moment où il devient malade) est de minimiser le risque de biais de sélection. Un cas prévalent, c'est-à-dire déjà malade avant d'être inclus dans l'étude est un cas incident ayant survécu et n'ayant pas été perdu de vue à la date d'inclusion dans l'étude. Si la maladie est grave, les cas prévalents inclus sont les patients les moins graves ayant survécu suffisamment longtemps pour participer à l'étude cas-témoins (biais de sélection par le pronostic). Si la maladie est potentiellement guérissable, les cas prévalents inclus sont ceux n'ayant pas guéri de leur maladie (biais de sélection par le recours au soin). Ces deux mécanismes de biais de sélection n'existent pas dans une étude réalisée sur des cas incidents (réponse D vraie). Un autre avantage est que la chronologie exposition – maladie est plus convaincante (diminution du risque de biais protopathique qui survient quand il est difficile d'établir que l'exposition étudiée a bien précédé la survenue de la maladie) quand les sujets peuvent être interrogés sur leur exposition passée au moment ou juste après le diagnostic de



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

leur maladie (réponse B vraie). La séquence chronologique exposition/maladie est toujours potentiellement litigieuse dans les études cas-témoins, mais encore plus si l'on utilise les cas prévalents. En revanche, la durée de la période d'inclusion dépend du rythme d'apparition des nouveaux cas, ce qui rallonge en général cette durée par rapport à une étude sur cas prévalents (c'est-à-dire une étude pouvant potentiellement inclure l'ensemble des malades présents à un moment donné) (réponse A fausse). Les deux derniers items (clause d'ignorance et exclusion secondaires) concernent les essais randomisés et non les études cas-témoins, et sont donc évidemment faux (réponses C et E fausses).

### Question 5

**Réponse exacte : B**

*Commentaires* Les cas et les témoins sont appariés sur l'âge et le sexe, leurs caractéristiques sont donc identiques sur ces deux critères.

La stratification correspond à considérer des sous-groupes (aussi appelés strates) définis en fonction des valeurs d'une (ou plusieurs) variable(s). La stratification aboutie rarement à des caractéristiques parfaitement identiques entre les deux groupes comparés pour chacune des strates définies par les variables de stratification. Elle est donc moins précise que l'appariement qui consiste à regrouper deux à deux des sujets (par exemple de cas et de témoins) partageants des caractéristiques identiques. Enfin, l'ajustement est une méthode d'analyse statistique appliquée une fois les sujets inclus et les données collectées, et qui n'a donc pas pour objectif d'équilibrer les cas et témoins au moment de leur inclusion dans l'étude.

### Question 6

**Réponse exacte : C**

*Commentaires* L'appariement est une méthode de prise en compte des facteurs de confusion, comme le sont la restriction et la stratification (permettant de prendre en compte des facteurs de confusion grâce au design de l'étude) ou l'ajustement (permettant de prendre en compte des facteurs de confusion à l'étape de l'analyse de l'étude). Si l'âge et le sexe sont des facteurs de confusion, l'appariement sur ces caractéristiques diminue le risque de biais de confusion (réponse C vraie) au moins sur ces facteurs. Le biais de mémoire est un biais de classement différentiel, lié au fait que l'information sur l'exposition est recueillie différemment entre les cas et les témoins. Par exemple, les individus qui ont été malades ont tendance à rechercher les causes possibles de la maladie. Ce risque de biais peut être minimisé si le niveau d'exposition peut être mesuré à l'aide d'une mesure objective (par exemple un dosage biologique), ou en organisant un recueil d'informations aussi semblable que possible chez les cas et les témoins (réponse A fausse). Le biais d'avance au diagnostic (augmentation apparente de la durée de survie des patients du fait d'un dépistage plus précoce de leur maladie), le biais d'indication (lié au fait que la probabilité qu'une intervention soit prescrite dépende du pronostic plus ou moins grave des sujets) et le biais protopathique (déjà abordé à la question 4) ne sont pas non plus pris en compte par cet appariement (réponses B, D et E fausses).

### Question 7

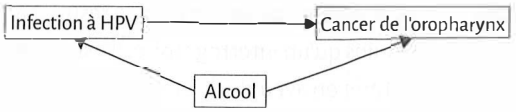
**Réponses exactes : B – D**

*Commentaires* La consommation d'alcool est un facteur de risque causal établi du cancer de l'oropharynx, et a potentiellement une influence sur le risque d'infection à HPV : il s'agit donc d'un facteur de confusion potentiel, qu'un appariement sur le statut alcoolique des sujets aurait permis de prendre en compte (réponses B et D vraies).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Schéma résumant les liens probables entre alcool, infection à HPV et cancer de l'oropharynx : l'alcool pourrait être un facteur de confusion dans l'association d'intérêt




A noter que les auteurs ont plutôt choisi de prendre en compte ce facteur de confusion potentiel durant l'analyse des données (ajustement de l'analyse sur ce facteur, avec, en plus, l'ajout d'une interaction avec l'exposition d'intérêt).

## Question 8

**Réponse exacte : C**

*Commentaires* Les pratiques sexuelles ne sont *a priori* pas liées de manière causale au risque de cancer de l'oropharynx, il ne peut donc s'agir d'un facteur de confusion (réponses D et E fausses). En effet, il s'agit plutôt d'un mode de contamination à HPV, mais pas de facteur de risque indépendant de la maladie (paragraphe 3 de la section Discussion : « Our data suggest that oral HPV infection is sexually acquired »). Si une association est observée avec le cancer de l'oropharynx dans cette étude (tableau 2), l'association passe donc par une augmentation du risque d'infection à HPV lié à ses pratiques. Apparier sur le comportement sexuel revient donc « un peu comme » à apparier sur le statut vis-à-vis de l'infection au HPV. A l'extrême, cela peut donc masquer une association causale réelle entre infection à HPV et cancer de l'oropharynx (réponse C vraie). Pour cette même raison, les auteurs n'ont pas ajusté l'analyse sur les comportements sexuels.

Schéma résumant les liens probable entre sexualité, infection à HPV et cancer de l'oropharynx : la sexualité n'est pas un facteur de confusion dans l'association d'intérêt




## Question 9

**Réponses exactes : C – D**

*Commentaires* Le fait que les femmes aient refusé plus souvent de participer à cette étude cas-témoins induit évidemment une surreprésentation des hommes dans l'échantillon analysé. Ce biais de sélection peut induire une diminution de la représentativité de l'échantillon (la population cible a probablement une proportion de femmes plus élevée), ce qui diminue la validité externe (ou la généralisabilité) de l'étude (réponses C et D vraies). L'appariement a pour objectif de contrôler des facteurs de confusion potentiels, afin d'augmenter la validité interne de l'étude (réponse A fausse). Dans cette étude, la proportion de sujets atteints d'un cancer de l'oropharynx est de 33 % par design (1 cas pour 2 témoins) (réponse B fausse).

## Question 10

**Réponses exactes : A – E**

*Commentaires* Pour limiter le biais de mémoire différentiel dans une étude cas-témoins, il est recommandé d'utiliser des questionnaires et des interviews standardisées avec des conditions de recueil comparables (ici auto administration du même questionnaire chez les cas et les témoins, deuxième paragraphe de la partie « data collection ») (réponse E correcte). De plus,



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

le statut vis-à-vis de l'infection à HPV ne repose pas sur l'interrogatoire des sujets inclus (au contraire du statut tabagique ou des pratiques sexuelles par exemple) mais sur une analyse sérologique en laboratoire, soit une mesure objective dont le résultat a moins de risque d'être influencé par le statut cas ou témoins qu'un interrogatoire. Il est également recommandé de mesurer le ou les expositions d'intérêt en aveugle du statut cas ou témoin. Aucune indication n'est donnée dans l'article sur ce point concernant les analyses sérologiques (réponse C fausse) ou les autres facteurs étudiés comme le statut tabagique ou alcoolique (il est toutefois probable que les cas aient répondu en connaissance de leur diagnostic de cancer, puisqu'ils étaient inclus dans l'étude suite à leur diagnostic). Enfin, le ou les expositions d'intérêt devraient être mesurées en aveugle de l'hypothèse de l'étude concernant l'association suspectée, mais là encore aucune information n'est rapportée dans l'article sur ce point (réponse B fausse). Les analyses multivariées permettent de prendre en compte des facteurs de confusion, mais pas le biais de mémoire (réponse D fausse).

Remarque sur le biais de mémoire dans les enquêtes cas-témoins.

Une étude de cohorte inclue des sujets au départ indemnes de maladie, et les suit en relevant les facteurs de risque d'intérêt jusqu'à l'apparition de la maladie (ou la fin de l'étude). Elle minimise donc le biais de mémoire différentiel entre les malades (« cas ») et les non malades (« témoins »), puisque l'information sur l'exposition est relevée avant l'apparition de la maladie. Une étude cas-témoin nichée dans une cohorte réalise une étude cas-témoin à partir des sujets malades (i. e. des cas apparus au cours du suivi de la cohorte) et non malades (i. e. témoins n'ayant pas présenté la maladie au cours du même suivi) d'une étude de cohorte. Elle profite donc des avantages d'une étude de cohorte concernant la prévention du biais de mémoire différentiel. Dans cet article, les cas sont issus d'une cohorte de patients atteints d'un cancer de l'oropharynx recrutés au moment du diagnostic de la maladie. Mais, les témoins ne sont pas issus de la même cohorte que les cas. Contrairement à ce que présentent les auteurs de l'article (« Our case-control study was nested within a longitudinal cohort study), l'étude ne répond donc pas à la définition habituelle d'une étude cas-témoins nichée dans une cohorte. Ce n'est en tout cas pas grâce à ce schéma d'étude que le biais de mémoire est minimisé.

### Question 11

**Réponse exacte : E**

*Commentaires* Dans une étude cas-témoins, la seule mesure d'association (mesurant le lien entre exposition et maladie) calculable est l'odds ratio.

### Question 12

**Réponse exacte : D**

*Commentaires* Un odds ratio (OR) (comme un risque relatif ou un hazard ratio) est toujours calculé par rapport à une classe de référence, par rapport à laquelle le « surrisque » de maladie (OR > 1) ou le « sous-risque » de maladie (OR < 1) est calculé. L'OR associé à la classe de référence est égal à 1 par construction.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 13**

**Réponse exacte : B**

*Commentaires* Aucun degré de signification n'est rapporté dans ce tableau. Ils auraient été tout à fait calculables, mais ils ne sont pas nécessaires pour conclure quant à la significativité des associations, les auteurs ayant rapportés les intervalles de confiance à 95 % autour des OR. Ces intervalles de confiance ne contiennent pas la valeur 1, donc les associations observées sont statistiquement significatives (et les degrés de signification inférieurs à 0,05).

**Question 14**

**Réponse exacte : E**

*Commentaires* Dans une étude cas-témoins, la prévalence de la maladie est fixée par le schéma de l'étude : elle est de 50 % quand l'étude inclut 1 témoin par cas, de 33 % quand l'étude inclut 2 témoins par cas, etc. La prévalence de la maladie dans l'échantillon de l'étude n'apporte strictement aucune information sur la prévalence de la maladie en population générale.

**Question 15**

**Réponse exacte : E**

*Commentaires* Une étude cas-témoins est de faible niveau de preuve, et il est globalement assez rare de prendre une décision de santé publique sur les seuls résultats d'une étude cas-témoins. Mais la raison principale pour laquelle aucune décision de santé publique concernant la vaccination ne peut être prise sur la base des résultats de cette seule étude est qu'il ne s'agit pas d'une étude évaluant la vaccination. Les études pouvant apporter des éléments sur cette stratégie pourraient aborder les questions suivantes : Quid de l'efficacité vaccinale ? Du nombre de cancer évités grâce à la vaccination ? Du rapport coût-efficacité ? etc.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 8 Étude cas-témoin

Mathieu Oriol

| | |
|---|---|
| **Titre de l'article** | **Radiotherapy- and Chemotherapy-Induced Myelodysplasia Syndrome: A Nationwide Population-Based Nested Case–Control Study** |
| **Auteurs** | Sun L-M, Lin C-L, Lin M-C, Liang J-A, Kao C-H |
| **Référence** | Medicine. 2015 ; 94(17) ) : e737 |
| **Lien vers l'article** | http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4603069/ |

**1** **Quel(s) est(sont) l'(les) élément(s) présent(s) dans l'introduction sur le(s) quel(s) les auteurs s'appuient pour le rationnel leur étude ?**

a. La combinaison radiothérapie et chimiothérapie a déjà été identifiée comme associée à un surrisque de développer un Syndrome MyéloDysplasique (MDS)
b. La prévalence des personnes vivant avec un cancer à Taïwan est en diminution depuis 2011
c. La proportion de MDS consécutives à un traitement est faible
d. Le traitement par radiothérapie a déjà été identifié comme facteur de risque de MDS dans plusieurs études interventionnelles
e. Les MDS sont des pathologies de mauvais pronostic

**2** **L'(les) objectif(s) de cette étude était(ent) de**

a. Démontrer que le traitement par radiothérapie ou chimiothérapie de certains cancers est un facteur de risque de MDS
b. Évaluer la proportion de patients ayant reçu un traitement par radiothérapie ou chimiothérapie à risque de décéder parmi le système NHI
c. Mettre en évidence un lien causal entre un antécédent de traitement par chimiothérapie et la survenue d'un MDS
d. Mettre en évidence une association entre un antécédent de traitement par radiothérapie et la survenue d'un MDS
e. Montrer que la radiothérapie est supérieure à la chimiothérapie dans le traitement des MDS

**3** **Quel est le type d'étude utilisé par les auteurs pour répondre à aux objectifs mentionnés à la fin de la partie Introduction ?**

a. Analytique
b. Cas-témoins
c. Cohorte
d. Interventionnelle
e. Observationnelle



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **Il s'agit d'une étude observationnelle analytique/étiologique de type cas-témoins nichée dans une cohorte**

a. L'OR peut être estimé dans les études cas-témoins
b. L'OR peut être estimé dans les études de cohorte
c. Le RR peut être estimé directement dans les études cas-témoins
d. Le RR peut être estimé directement dans les études cas-témoins nichées dans une cohorte
e. Le RR peut être estimé dans les études de cohorte

**5** **Quels sont les éléments présentés par les auteurs qui permettent de dire que la cohorte LHID2000 est représentative de la base de données NHRID ?**

a. La cohorte LHID2000 représente plus de 3 % de la base de données NHRID
b. Les caractéristiques des patients de la cohorte LHID2000 ne diffèrent pas de celles de la base de données NHRID
c. Les données de patients sont anonymisées
d. Les diagnostics sont standardisés à l'aide d'une classification internationale (ICD-9-CM)
e. Les individus de la cohorte LHID2000 sont tirés au sort

**6** **Le programme NHI rassemble les données médicales de la quasi-totalité de la population taïwanaise. Quel(s) est(sont) le(s) biais qui est(sont) minimisé(s) avec cette approche ?**

a. Attrition
b. Classement
c. Confusion
d. Mesure/information
e. Sélection

**7** **La standardisation du diagnostic de cancer à partir de la classification ICD-9-CM permet de minimiser le(s) biais de**

a. Attrition
b. Classement
c. Confusion
d. Mesure/information
e. Sélection

**8** **Quelles sont les caractéristiques des cas ?**

a. Le diagnostic pouvait être confirmé de manière anatomopathologique
b. Patient ayant présenté un nouveau cancer entre 2000 et 2011
c. Patient ayant présenté un MDS après 2011
d. Patient ayant présenté une leucémie myéloïde chronique entre 2000 et 2011
e. Patient n'ayant pas présenté de MDS



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**9** **Les auteurs ont apparié les cas avec les témoins sur plusieurs caractéristiques afin de minimiser le(s) biais de**

a. Attrition
b. Classement
c. Confusion
d. Mesure/information
e. Sélection

**10** **Dans le tableau 1, pour comparer les caractéristiques des cas et des témoins, les auteurs ont utilisé**

a. Des analyses descriptives
b. Des analyses multivariées
c. Des analyses univariées
d. Des tests de comparaison de pourcentages
e. Des tests paramétriques

**11** **À partir du tableau 1, il est possible d'affirmer que**

a. Chez les témoins, la proportion de patients traités par antimétabolites est statistiquement plus importante que celle des patients traités par agents alkylants
b. L'absence de différence significative sur les variables « sexe » et « âge » montre que l'appariement entre les cas et les témoins a été bien réalisé
c. La proportion de femmes chez les cas est statistiquement moins importante que chez les témoins
d. Le pourcentage de patients traités par radiothérapie statistiquement plus important chez les cas est à l'origine d'un biais de confusion
e. Les patients diabétiques sont statistiquement plus nombreux chez les cas que chez les témoins

**12** **Dans le Tableau 2**

a. L'OR brut significatif pour le diabète était prévisible car la proportion de diabétiques était significativement différente dans le tableau 1
b. La variable « Chronic obstructive pulmonary disease » est un facteur de confusion
c. La proportion d'AVC (stroke) est significativement plus importante chez les cas par rapport aux témoins
d. La radiothérapie est un facteur de risque indépendant de MDS
e. Les antimétabolites sont un facteur protecteur indépendant de MDS



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**13** **Dans le tableau 3**

a. Chez les patients ayant reçu une chimiothérapie pour un cancer du col de l'utérus (cervix), le risque de MDS était significativement augmenté par rapport aux patients sans chimiothérapie
b. L'OR 1.00 (Reference) est une comparaison des cas et des témoins
c. La chimiothérapie est un facteur de risque indépendant de cancer de l'estomac
d. La radiothérapie est un facteur de risque indépendant de cancer colorectal
e. Le très faible nombre de patients atteints de leucémie lymphoblastique est responsable de la très grande largeur de l'intervalle de confiance de l'OR correspondant : OR = 2,88 [0,02-339,9]

**14** **Quel(s) est(sont) le(s) critère(s) de Hill qui peut(peuvent) être relevé(s) dans cette étude ?**

a. Force de l'association
b. Relation dose-effet
c. Reproductibilité de l'association
d. Plausibilité de l'association
e. Temporalité de l'association

**15** **Les résultats de cette étude permettent**

a. De mettre en évidence un lien de cause à effet entre un antécédent de traitement par chimiothérapie chez les patients atteints de cancer et la survenue d'un MDS : OR = 1,51 [1,25-1,82], $p < 0,001$
b. De mettre en évidence une association statistiquement significative entre la chimiothérapie et la survenue d'un cancer du poumon : OR = 2,67 [1,07-6,67], $p < 0,05$
c. De mettre en évidence une association statistiquement significative entre un traitement par radiothérapie et la survenue d'un MDS chez les patients atteints d'un cancer du rein : OR = 5,59 [1,36-23,1], $p < 0,05$
d. De préconiser l'arrêt des traitements par chimiothérapie chez les patientes atteintes de cancer de l'utérus/endomètre car ce traitement est identifié comme facteur de risque de MDS : OR = 7,59 [1,07-53,6], $p < 0,05$
e. De préconiser la poursuite des traitements par radiothérapie chez les patients atteints de cancer de l'œsophage car ce traitement n'est pas identifié comme facteur de risque de MDS : OR = 0,84 [0,21-3,33], $p > 0,05$



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 1

**Réponses exactes : C – E**

*Commentaires*

A. La combinaison n'a pas été identifiée. Par contre la chimiothérapie ou la radiothérapie seules ont été reliées à un surrisque de MDS.

B. « The proportion of long-term cancer survivors is rising », donc augmentation de la prévalence.

C. « Although the absolute number of treatment-related MDS cases is small ».

D. Aucune mention n'est faite à des études interventionnelles dans l'introduction.

E. « The poor prognosis of MDS warrants concern ».

### Question 2

**Réponses exactes : A – D**

*Commentaires*

A. L'étude étant observationnelle, il est possible de mettre en évidence des facteurs de risques d'une maladie.

B. Il ne s'agit pas d'une étude mesurant la survie des patients traités par radiothérapie ou chimiothérapie.

C. L'étude n'étant pas interventionnelle, il n'est pas possible de mettre en évidence un lien de cause à effet.

D. Létude étant observationnelle, il est possible de mettre en évidence une association entre une exposition et une maladie.

E. Il ne s'agit pas d'une étude interventionnelle, seule type d'étude permettant de tester la supériorité d'un traitement par rapport à un autre en limitant les biais, notamment d'indication.

### Question 3

**Réponses exactes : A – B – E**

*Commentaires* Il s'agit d'une étude observationnelle analytique/étiologique de type cas-témoins nichée dans une cohorte. Il ne s'agit pas d'une étude de cohorte. L'étude de cohorte sous-jacente sert juste de base de données pour sélectionner les cas et les témoins.

### Question 4

**Réponses exactes : A – B – E**

*Commentaires* Le RR ne peut être estimé dans les études cas-témoins car la proportion malades/non-malades est fixée par l'investigateur (ici 1 cas pour 4 témoins) et ne correspond pas à la répartition que l'on peut trouver en population générale ou dans la cohorte dont sont issus les patients pour cette étude.

### Question 5

**Réponses exactes : B – E**

*Commentaires*

A. Le simple pourcentage d'individus inclus dans l'échantillon ne conditionne pas sa représentativité.

B. « No statistically significant differences ».



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

C. C'est un plus en termes de protection des données des patients mais cela n'intervient pas dans la représentativité de l'échantillon.

D. Cela permet de minimiser le biais de classement mais cela n'intervient pas dans la représentativité de l'échantillon.

E. « Randomly sampled ».

**Question 6**

**Réponse exacte : E**

**Indispensable : E**

*Commentaires* Par cette approche, les auteurs sont sûrs d'avoir la totalité de la population taïwanaise et de ne pas sélectionner artificiellement une partie de leur population.

**Question 7**

**Réponses exactes : B – D**

*Commentaires* Les cancers sont tous classés de la même façon via la classification ICD-9-CM. La standardisation permet de minimiser le biais de mesure/information qui est un biais de classement.

**Question 8**

**Réponses exactes : A – B**

*Commentaires*

A. « Confirmatory pathological ».

B. C'est une des caractéristiques du Registry for Catastrophic Illness Database qui valide les cas de cancer.

C. Ce sont les cas de MDS de 2000 à 2011 qui sont inclus.

D. La leucémie myéloïde chronique est un critère d'exclusion.

E. Ce sont les témoins qui ne présentent pas de MDS.

**Question 9**

**Réponse exacte : C**

**Indispensable : C**

*Commentaires* Le but de l'appariement est de rendre comparables deux groupes en termes de facteurs de confusion potentiels.

**Question 10**

**Réponses exactes : C – D – E**

*Commentaires* Les auteurs ont utilisé le test du Chi-2 et le test t de Student, tests paramétriques, pour comparer les variables 2 à 2. Ce sont des analyses univariées. On parle parfois d'analyses bivariées pour souligner que dans ces analyses on regarde la différence de distribution d'une variable en fonction d'une autre (i. e. on « croise » deux variables)



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 11

**Réponses exactes : B – E**

*Commentaires*

A. La comparaison se fait ici entre cas et témoins et non entre les caractéristiques des cas.

B. Il ne doit pas être retrouvé de différence significative au niveau des variables d'appariement. Sinon l'étude présente un potentiel biais de confusion.

C. 50,8 % chez les cas *versus* 50,9 % chez les témoins mais la différence n'est pas statistiquement significative.

D. La radiothérapie est un des « facteurs à expliquer » étudiés de survenue de MDS.

E. 20,1 % chez les cas *versus* 16,8 % chez les témoins, p = 0,006 donc la différence est statistiquement significative.

### Question 12

**Réponses exactes : A – B – C – D**

*Commentaires*

A. Ce sont les mêmes données exprimées de 2 façons différentes.

B. La variable BPCO était significative en analyse bivariée et devient non significative en analyse multivariée.

C. L'OR est supérieur à 1 et ne contient pas 1.

D. OR = 1,53 [1,33-1,77], p < 0,001.

E. L'OR contient 1 : OR = 0,91 [0,77-1,08], non significatif.

### Question 13

**Réponses exactes : A – E**

*Commentaires*

A. OR = 2,41 [1,16-5,00], p < 0,05

B. Il s'agit du risque de présenter, pour chaque type de cancer, un MDS chez les patients n'ayant pas été traités par radiothérapie par rapport aux patients n'ayant pas été traités par radiothérapie, c'est-à-dire eux-mêmes. Cela constitue donc la catégorie de référence par rapport à laquelle vont être comparés les patients ayant été traités par radiothérapie. Le même raisonnement peut être appliqué pour la chimiothérapie.

C. C'est le risque de MDS qui est étudié ici chez les patients atteints de cancer de l'estomac.

D. C'est le risque de MDS qui est étudié ici chez les patients atteints de cancer colorectal.

E. Oui car plus l'effectif est important plus l'intervalle de confiance est petit et proche de l'OR de la population source.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 14

**Réponses exactes : A – C – D – E**

*Commentaires*

A. Dans le tableau 5, on peut retrouver un OR significatif à 37 dans le cancer de l'utérus pour la variable radiothérapie chimiothérapie *versus* absence de traitement.

B. Les auteurs expliquent dans la discussion que les doses de chimiothérapie et radiothérapie n'étaient pas disponibles dans la base du NHI.

C. Les auteurs expliquent dans la discussion que certains résultats concordent avec ceux retrouvés dans d'autres études, par exemple pour le cancer du sein.

D. Les auteurs expliquent dans la discussion que certains effets ont été retrouvés au niveau chromosomique, par exemple pour les agents alkylants.

E. Le fait que les cas et les témoins soient issus d'une cohorte permet d'être sûr de la temporalité. Les auteurs précisent d'ailleurs que les patients ayant présenté un MDS avant leur diagnostic de cancer ont été exclus.

## Question 15

**Réponse exacte : C**

**Indispensable : C**

*Commentaires*

A. Il n'est pas possible de mettre en évidence de lien de cause à effet dans les études observationnelles.

B. La maladie étudiée dans cette étude est le MDS et non le cancer du poumon.

C. L'OR est significatif.

D. Le bénéfice d'un traitement par chimiothérapie chez les patientes atteintes de cancer de l'utérus/endomètre (i. e. décès de la patiente si absence de traitement) est jugé supérieur au risque de développer un MDS après cancer.

E. L'absence d'association entre le traitement par radiothérapie et la survenue d'un MDS n'entre pas en compte dans la décision de traiter un patient atteint d'un cancer de l'œsophage, *cf.* réponse ci-dessus.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 9 Étude descriptive (transversale)

Paul de Boissieu

| | |
|---|---|
| **Titre de l'article** | **Venous thromboembolism risk and prophylaxis in the acute hospital care setting (ENDORSE study): a multinational cross-sectional study** |
| **Auteurs** | Cohen AT, Tapson VF, Bergmann JF, Goldhaber SZ, Kakkar AK, Deslandes B, Huang W, Zayaruzny M, Emery L, Anderson FA |
| **Référence** | Lancet 2008 ; 371(9610):387-94 |
| **Lien vers l'article** | http://dx.doi.org/ 10.1016/S0140-6736(08)60202-0 |

**1** **D'après vous, les objectifs de cette étude étaient**

a. D'estimer le nombre de patients à risque de maladie thromboembolique veineuse
b. D'évaluer l'efficacité de la prophylaxie médicamenteuse sur la survenue de maladie thromboembolique veineuse
c. De comparer l'efficacité de la prophylaxie anti- maladie thromboembolique veineuse entre les patients pris en charge en médecine et ceux pris en charge en chirurgie
d. D'évaluer le nombre de patients hospitalisés en médecine et en chirurgie, dans les services d'urgences
e. De déterminer, parmi les patients à risque de maladie thromboembolique veineuse, la proportion recevant une prophylaxie conforme aux recommandations

**2** **De quel type d'étude s'agit-il ?**

a. Essai randomisé contrôlé
b. Étude observationnelle
c. Étude cas-témoin
d. Étude de cohorte
e. Étude interventionnelle

**3** **Parmi les pays participants à cette étude, lesquels ont constitué un échantillon représentatif de leurs hôpitaux ?**

a. USA
b. France
c. Brésil
d. Allemagne
e. Russie



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **Parmi les éléments suivants, lesquels étaient des critères d'inclusion concernant les patients ?**

a. Patient traité pour une maladie thromboembolique active
b. Patient hospitalisé dans un service éligible
c. Patient présentant un risque de maladie thromboembolique
d. Patient présentant un risque élevé de maladie thromboembolique
e. Patient de 18 ans ou plus (chirurgie) ou de 40 ans ou plus (médecine)

**5** **Concernant le résultat suivant : « 35 329 (51,8 % ; 95 % -CI = 51,4-52,2 % ; range = 35,6-72,6 %) of the enrolled patients were deemed to be at risk for VTE », quelles sont les propositions justes ?**

a. Il s'agit d'un résultat principal
b. Il s'agit d'un résultat secondaire
c. La proportion de patients à risque de maladie thromboembolique veineuse variait de 35,6 % à 72,6 % selon les pays
d. La proportion de patients à risque de maladie thromboembolique veineuse variait de 51,4 % à 52,2 % selon les pays
e. Il manque le petit p de l'intervalle de confiance à 95 % pour savoir si la proportion observée est statistiquement significative

**6** **Quelle est la proposition juste concernant la prescription d'une anticoagulation chez un patient qui présente une hémorragie cérébrale ?**

a. Un médecin ne peut pas prescrire un médicament dès lors que le patient présente une contre-indication dûment mentionnée sur la notice
b. En cas de plainte, la responsabilité du médecin sera engagée uniquement si le patient présentait des contre-indications à la prescription
c. Si l'anticoagulation est prescrite malgré la présence d'une contre-indication, le patient sera moins bien remboursé
d. Les médecins ont l'obligation de suivre les recommandations de prophylaxie chez les patients à risque
e. La responsabilité du médecin peut être engagée si le patient décède suite à cette prescription

**7** **L'anticoagulation médicamenteuse par héparine de bas-poids moléculaire (HBPM) était la forme la plus fréquente de prophylaxie utilisée dans cette étude. Concernant les HBPM, quelles sont les propositions justes ?**

a. Leur efficacité est surveillée grâce à la réalisation régulière d'INR
b. Elles ont une activité anti Xa
c. Suite à l'initiation du traitement, la survenue d'une thrombopénie est classique et bénigne
d. Elles sont administrées par voie sous-cutanée
e. Elles peuvent être utilisées à visée préventive et curative



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**8** **Concernant l'utilisation d'une prophylaxie conforme aux recommandations « 11 613 (58,5 % ; 95 % -CI = 57,8-59,2) [...] were surgical and 6119 (39,5 % ; 95 % -CI = 38,7-40,3) [...] were medical patients », quelles sont les propositions justes ?**

a. La proportion de patients recevant une prophylaxie conforme aux recommandations était plus élevée chez les patients chirurgicaux
b. La proportion de patients recevant une prophylaxie conforme aux recommandations était significativement plus élevée chez les patients chirurgicaux
c. Entre 2006 et 2007, l'utilisation conforme d'une prophylaxie a varié entre 57,8 % et 59,2 % chez 95 % des patients de chirurgie
d. Entre 2006 et 2007, l'utilisation conforme d'une prophylaxie a varié entre 38,7 % et 40,3 % chez 95 % des patients de médecine
e. Entre 2006 et 2007, l'utilisation conforme d'une prophylaxie a été estimée à 58,5 % chez les patients hospitalisés en chirurgie, à risque de maladie thromboembolique veineuse

**9** **À propos du tableau 5, présentant le détail des résultats par pays, quelles sont les propositions justes ?**

a. Pour la France, 1 927 patients au total ont été évalués
b. Pour la France, 2 844 patients au total ont été évalués
c. Pour la France, 36 % de l'ensemble des patients hospitalisés en service de médecine étaient à risque de maladie thromboembolique veineuse
d. Pour la France, 78 % de l'ensemble des patients hospitalisés en service de chirurgie étaient à risque de maladie thromboembolique veineuse
e. Ce tableau a été réalisé à partir des 68 183 patients inclus dans cette étude

**10** **Concernant la réalisation de cette étude, quelles sont les propositions justes ?**

a. Cette étude a été financée (sponsorisée) par l'industrie pharmaceutique
b. Cette étude a été financée (sponsorisée) par le Ministère de la Santé
c. L'industrie pharmaceutique a fourni une aide lors de la rédaction de cet article
d. Le Ministère de la Santé a fourni une aide lors de la rédaction de cet article
e. Une étude ne peut pas être réalisée en France sans un financement du Ministère de la Santé

**11** **Concernant la phrase suivante : « a minimum of 450 patients per analysis group are required » (statistical analysis), quelle(s) est(sont) la(les) proposition(s) juste(s) ?**

a. 450 patients par groupe représente le nombre de sujets nécessaires
b. 450 patients par groupe représente la puissance de l'étude
c. En dessous de 450 patients par groupe, le résultat ne sera pas représentatif
d. En dessous de 450 patients par groupe, le résultat ne sera pas significatif
e. Toutes les réponses précédentes sont fausses



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**12** **Concernant la phrase suivante « Of the population at risk for VTE, 1549 (10 %) medical patients and 1 791 (9 %) surgical patients were considered to have a high bleeding risk, sufficient to present a contraindication to anticoagulation prophylaxis », quelles sont les propositions justes ?**

a. L'inclusion de ces sujets, pour calculer la proportion de sujets à risques recevant une prophylaxie adaptée, est justifiée
b. Ces sujets n'auraient pas dû être inclus dans le calcul de la proportion de sujets à risque recevant une prophylaxie adaptée
c. Il s'agit de patients ayant présentés un effet indésirable médicamenteux
d. L'inclusion de ces sujets sous-estime la proportion de patients recevant une prophylaxie adaptée
e. Il s'agit de patients ayant des critères d'exclusion

**13** **Selon la Haute Autorité de Santé (HAS), on peut dire de cette étude qu'elle est de**

a. Niveau de preuve 1
b. Niveau de preuve 4
c. Grade de recommandation A
d. Grade de recommandation B
e. Grade de recommandation C

**14** **Concernant la partie « Conflict of interest statement » (fin de l'article), quelles sont les propositions justes ?**

a. En France, la déclaration de lien d'intérêt entre une industrie de santé et un médecin est obligatoire
b. En France, la déclaration de lien d'intérêt entre une industrie de santé et un médecin est publique
c. En France, la déclaration de lien d'intérêt entre une industrie de santé et un médecin est confidentielle
d. Un lien d'intérêt peut impacter la puissance de cette étude
e. Un lien d'intérêt peut impacter l'objectivité des auteurs de cette étude

**15** **Parmi les limites citées par les auteurs dans la partie « Discussions », quelles sont celles pouvant impacter la validité interne de l'étude ?**

a. La connaissance par les médecins de l'existence de cette étude
b. La non-représentativité
c. L'utilisation des dossiers médicaux, pas toujours mis à jour
d. Erreur lors de la saisie des données
e. La faible proportion de traumatisés graves dans cette étude



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 1

**Réponses exactes : A – E**

*Commentaires* Dans la plupart des articles scientifiques, le ou les objectifs principaux sont donnés dans la dernière phrase (ou le dernier paragraphe) de l'introduction. C'est le cas dans cet article : « to assess the number of patients at risk for VTE in the acute care hospital setting and to determine the proportion of these at-risk patients who received prophylaxis as recommended by the... ». Attention, même si le résumé présente parfois l'objectif principal de l'article, il est préférable de se référer au texte entier pour être sûr de ne pas manquer une partie de l'objectif.

## Question 2

**Réponse exacte : B**

*Commentaires* Cette question renvoie aux différents types d'études. Il existe deux types distincts d'étude : interventionnelle ou observationnelle. Dès lors que la prise en charge habituelle du patient est modifiée (traitement, posologie, examens supplémentaires...), il s'agit d'une étude interventionnelle. Si rien n'est modifiée (on se contente d'observer/décrire sans intervenir), alors l'étude est observationnelle.

A et E : Pour cette étude, la partie « Méthode » ne décrit aucun changement dans la prise en charge habituelle des patients. Aucun test, aucun examen supplémentaire n'est imposé par l'étude. Il s'agit donc d'une étude observationnelle. L'essai randomisé étant une étude interventionnelle, les réponses A et E sont fausses.

B : Parmi les études observationnelles, il existe plusieurs types d'étude : cas-témoin, cohorte, transversale... Le dernier paragraphe de l'introduction donne précisément le type d'étude réalisée : « observational, cross-sectional ». Il s'agit d'une étude observationnelle de type transversale. Le but de ce type d'étude est de permettre un état des lieux descriptif.

C : Lors d'une étude cas-témoin, la population de l'étude est composée de patients présentant la maladie d'intérêt (les cas) et de patients ne présentant pas la maladie d'intérêt (les témoins). Ils sont d'ailleurs sélectionnés sur ce critère (avoir ou non la pathologie d'intérêt). Dans la partie « Méthode », à la sous-section « Patients », il est précisé que les patients sont sélectionnés à partir des services hospitaliers « medical wards [...] surgical wards » et non pas sur selon leur statut maladie thromboembolique. Ce n'est donc pas un cas-témoin.

D : L'étude de cohorte est une étude observationnelle durant laquelle des patients, présentant une ou plusieurs caractéristiques communes (année de naissance, lieu d'hospitalisation, début d'un traitement X...), sont suivis dans le temps. Pour cette étude, la partie « Méthode » ne mentionne pas de suivi des patients dans le temps. Il ne s'agit donc pas d'une étude de cohorte.

## Question 3

**Réponses exactes : A – B – D**

*Commentaires* La sélection des hôpitaux inclus dans cette étude est décrite dans la section « Méthode », dans le chapitre « Procédures ». Il y est précisé que les hôpitaux participants à cette étude, ont été tirés au sort (« selected at random ») dans les 32 pays participants. Il s'agit donc d'échantillons représentatifs (car tirés au sort). Néanmoins, il est apporté une précision au paragraphe suivant, qui va altérer la représentativité de certains hôpitaux. Une liste de pays est mentionnée, pour lesquels les hôpitaux initialement tirés au sort n'ont pas été jugés aptes à être inclus dans cette étude (expérience de recherche trop limitée, ou capacité de recueil de données insuffisante). En conséquence, pour ces pays, les hôpitaux ont été sélectionnés individuellement, par les responsables de l'étude. Ces hôpitaux n'étant plus tirés au sort, la



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

représentativité n'est plus garantie. Le Brésil et la Russie font partie des pays pour laquelle une sélection individuelle a été mise en place. Les USA, la France et l'Allemagne ne figurent pas dans cette liste, et ont donc inclus un échantillon représentatif de leurs hôpitaux.

■ **Question 4**
**Réponses exactes : B – E**
**Inadmissibles : C – D**

*Commentaires* Il est nécessaire de bien comprendre les chapitres « Procédures » et « Patient » de la section « Méthode » pour répondre à cette question. Dans un premier temps, ce sont des hôpitaux qui ont été tirés au sort dans chaque pays (avec la petite subtilité vu à la question 3). Pour être tirés au sort, les hôpitaux devaient remplir un certain nombre de critères (qu'on pourrait qualifier de critère d'inclusion pour les hôpitaux). Ces critères sont détaillés dans le premier paragraphe du chapitre « Procédures » : « if they contained more than 50 beds... ». Dans un second temps, seules certaines unités d'hospitalisation des hôpitaux tirés au sort étaient ensuite utilisées pour cette étude. D'autres unités étaient exclues : « Hospitals wards were eligible if they... [...] and palliative care wards were excluded ». Cette fois, les unités n'étaient pas tirées au sort : toutes les unités éligibles étaient utilisées : « All eligible wards within enrolled hospitals were included ». Dans un troisième (et dernier) temps, les patients ont été inclus. Il s'agit du processus détaillé au chapitre « Patients ». C'est sur cette partie que porte la question.

A : L'un des objectifs de cette étude est d'évaluer, chez les patients à risque de maladie thromboembolique veineuse, la proportion de ceux traités par prophylaxie. Les patients inclus dans cette étude pouvaient donc être traités ou non-traités. La présence d'un traitement n'était pas un critère d'inclusion.

B : Au début du chapitre « Patients », il est indiqué que les patients hospitalisés dans les services éligibles sont inclus

C et D : L'un des objectifs de cette étude est d'estimer le nombre de patients à risque de maladie thromboembolique veineuse. Ce risque doit donc être évalué chez tous les patients inclus : « Enrolled patients were assessed for risk for VTE in accordance with the 2004 ACCP guidelines ». Les patients pouvaient donc avoir, ou non, un risque de maladie thromboembolique veineuse. Ne sélectionner que des patients à risque (ou à fort risque) de maladie thromboembolique veineuse, rendrait impossible l'évaluation d'une proportion de patients à risque.

E : L'âge des patients, selon le type d'unité (médicale ou chirurgicale) est mentionné comme critère d'inclusion dès le début du chapitre « Patients ».

■ **Question 5**
**Réponses exactes : A – C**
**Inadmissible : E**

*Commentaires* Cette question traite de l'interprétation d'un résultat principal de cette étude.

A et B : Ce résultat présente le nombre (et la proportion correspondante) de patients présentant un risque de maladie thromboembolique veineuse. Comme mentionné à la fin de l'introduction, il s'agit d'un objectif principal de cette étude (l'autre étant la proportion de patient bénéficiant d'une prophylaxie).

C et D : Ces deux propositions portent sur l'interprétation d'un intervalle de confiance à 95 % (« confidence interval » en anglais : CI) et d'une étendue (« range » en anglais). L'étendue (ou l'amplitude) donne une indication sur les valeurs extrêmes, que sont le minimum et la maximum. Il est précisé dans le chapitre « Statistical analysis » que « le « between-country ranges were



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

calculated ». Cela signifie donc que le « range » présenté ici, concerne les différences selon les pays. La proportion de patients à risque de maladie thromboembolique veineuse variait donc de 35,6 % à 72,6 % entre les pays. L'intervalle de confiance à 95 % ($IC_{95\ \%}$) est représenté par deux valeurs/bornes : une inférieure et une supérieure. Dans 95 % des cas/des estimations, la « vraie » valeur de ce que l'on cherche à estimer (ici la proportion de patients à risque de maladie thromboembolique veineuse) se situera entre les bornes de cet intervalle. Donc dans 95 % des cas, la proportion vraie de patients à risque de maladie thromboembolique veineuse sera comprise entre 51,4 % et 52,2 %. Pour cette étude, la proportion estimée était de 51,8 %.

E : Les seuls cas où le petit p peut apporter de l'information supplémentaire à l'interprétation d'un intervalle de confiance concernent l'évaluation de risque relatif ou d'odds ratio. Une proportion observée est une donnée brute : sur cet échantillon, 51,8 % des patients étaient à risques de maladie thromboembolique veineuse. Il n'y a pas besoin de petit p pour interpréter une donnée observationnelle seule.

## Question 6

**Réponse exacte : E**

*Commentaires*

A : Les médicaments, pour être commercialisés (prescrits, puis administrés) doivent obtenir une autorisation de mise sur le marché (AMM). Dans cette AMM, figurent notamment la ou les indications (ce pour quoi le médicament peut être prescrit) ainsi que les contre-indications.

En principe (mais ce n'est pas une obligation), la prescription doit être conforme à l'AMM. Le médecin reste libre de suivre ou non ces indications, ainsi que de respecter ou non les contre-indications.

Bien évidemment, la prescription d'une anticoagulation lors d'une hémorragie cérébrale (ou de tout autre médicament en présence d'une contre-indication) doit être impérativement justifiée par un rapport bénéfice-risque favorable.

B : La responsabilité du médecin est engagée dès lors qu'il prescrit un médicament. Que le patient présente ou non une contre-indication, sa responsabilité sera engagée en cas de plainte du patient. Tous les médicaments peuvent entraîner des effets indésirables, potentiellement graves, même en l'absence de contre-indication. L'information du patient sur ces risques potentiels, ainsi que la déclaration de pharmacovigilance en cas d'effet indésirables, sont des obligations légales.

C : En France, le remboursement d'un médicament ne dépend pas de son indication. L'anticoagulation sera remboursée de la même manière qu'elle soit utilisée en préventif ou en curatif, y compris en présence d'une contre-indication.

D : Comme leur nom l'indique, il s'agit de recommandation. Bien qu'elles représentent la position de sociétés académiques, ces recommandations n'ont pas de valeur contraignante.

E : En cas de décès, toutes les responsabilités du médecin peuvent être engagées : civile (réparation), pénale (faute intentionnelle) et ordinale (faute déontologique).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 7

**Réponses exactes : B – D – E**

**Inadmissible : C**

*Commentaires*

A : L'INR (pour International Normalized Ratio) est utilisé pour le suivi des patients traités par anti-vitamine K (AVK), et non pas pour le suivi des HBPM.

C : la survenue d'une thrombopénie dans les suites de l'initiation d'un traitement par HBPM doit faire suspecter une thrombopénie induite par l'héparine (TIH). La TIH, bien que rare, reste une pathologie grave, nécessitant une prise en charge en urgence.

## Question 8

**Réponses exactes : A – E**

*Commentaires*

A et B : Classiquement, les études observationnelles transversales sont réalisées afin de décrire un phénomène. Dans le cas des articles médicaux, elles sont utilisées pour décrire un état de santé, une prise en charge, un système de soin... Leur but est essentiellement descriptif, et repose sur des pourcentages ou des moyennes. En général, aucun test statistique n'y est réalisé. Dans ce résultat, les taux de prophylaxie conforme des patients médicaux et chirurgicaux sont présentés. Aucune comparaison n'est effectuée : aucun test statistique n'est mentionné dans le chapitre « Statistical analysis » et aucun « p » n'est donné. On peut simplement dire ce que l'on observe : 58,5 % > 39,5 %.

C, D et E : ces réponses renvoient à la définition même de l'intervalle de confiance à 95 %, ainsi qu'à la définition de la période d'inclusion dans une étude transversale. L'intervalle de confiance d'une proportion (*cf.* question 5 pour la définition) est fixe pour une étude donnée. Il est calculé une fois pour toutes, et ne varie pas.

Une étude transversale représente une photographie, à un moment donné, d'un phénomène de santé. On s'attend donc à voir une date fixe : « voici le nombre de patient à risque au 5 novembre 2007 ». Il peut donc sembler déroutant de constater que les inclusions ont eu lieu entre 2006 et 2007. Comme déjà abordé, les études transversales (à l'inverse des études de cohorte) ne suivent pas les patients dans le temps. Ici, le chapitre « Procedures » nous apporte des précisions sur cette période d'inclusion : « because the typical hospital was too large to allow [...] all patients enrolments in 1 day, investigators were asked to organize data abstraction to complete at least one floor or ward on any particular day ». C'est donc la logistique de cette étude qui explique que le recueil des données se soit étendu sur plusieurs mois : il était impossible de tout réaliser le jour même, à cause de la taille de certains hôpitaux. De plus, certains pays ayant des obligations réglementaires particulières (consentement signé), cela a également du contribuer à un recueil décalé dans ces derniers, à cause du temps nécessaire à la réalisation de ces obligations.

La proportion de patient à risque en chirurgie n'a donc pas été suivie dans le temps (elle n'a pas variée entre 2006 et 2007, mais a été évaluée sur cette période). De plus, les réponses C et D ne correspondent pas à la définition d'un intervalle de confiance.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 9

**Réponses exactes : B – C – D – E**

*Commentaires*

A et B : Attention à ne pas lire trop rapidement les informations du tableau. En reprenant l'intitulé des colonnes en haut du tableau 5, on peut voir que les 1 927 patients ne sont que les patients hospitalisés en médecine. Pour avoir le total des patients évalués, il faut y ajouter les 917 patients hospitalisés en chirurgie.

C et D : Il s'agit là encore, de bien sélectionner la colonne d'intérêt pour lire le résultat correspondant. L'intitulé de ces deux réponses se rapporte aux colonnes « at-risk medical patients » et « at-risk surgical patients ».

E : La somme du total des patients (dernière ligne du tableau) hospitalisés en médecine (37356) et ceux hospitalisés en chirurgie (30 827) donne bien les 68 183 patients inclus dans cette étude.

## Question 10

**Réponses exactes : A – C**

*Commentaires*

A et B : En général, les informations concernant les éventuelles déclarations de conflits d'intérêts et le financement des études sont situées après la conclusion, juste avant les références. Pour cet article, il est clairement indiqué que « this study was sponsored by [Nom de l'industriel concerné] ». Il s'agit donc bien d'une étude financée par l'industrie pharmaceutique. En revanche, il n'est pas fait mention d'un quelconque financement par le Ministère de la Santé français.

C et D : Là encore, ces informations se situent dans la même partie de l'article. Il est là aussi indiqué que « editorial support for this article was provided by [Nom de l'industriel concerné] ». Ce qui signifie que l'industriel a bien apporté une aide (un support) lors de la rédaction de cet article. De même que pour le financement, le Ministère de la Santé n'a pas participé à la rédaction de cet article.

E : Pour réaliser une étude en France, il suffit de respecter les contraintes réglementaires (les lois) en vigueur. Ainsi, une étude peut être réalisée et financée à 100 % par l'industrie pharmaceutique, du moment qu'elle se conforme aux obligations réglementaires nécessaires.

## Question 11

**Réponse exacte : A**

*Commentaires* Dans cette étude transversale, l'objectif principal était d'évaluer le nombre de sujets à risque de maladie thromboembolique veineuse (dernière phrase de l'introduction). Il s'agit d'une variable qualitative binaire (à risque *versus* non à risque), qui sera exprimée en pourcentage. Le calcul du nombre de sujets nécessaires, dans une étude transversale, permet de s'assurer que la précision de l'estimation sera suffisante. Comment savoir si la mesure d'une variable qualitative binaire (un pourcentage) est précise ? Il faut regarder la largeur (l'amplitude) de l'intervalle de confiance. Plus l'intervalle est large, moins la mesure est précise (et inversement : plus l'intervalle est étroit, plus la mesure est précise). La largeur de l'intervalle de confiance va dépendre, entre autres, du nombre de sujets inclus. Plus le nombre de sujet est important, plus l'estimation sera précise.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

La lecture de la partie « statistical analysis » nous permet de comprendre que les auteurs souhaitaient avoir une précision de 4 % (*cf.* intervalle de confiance) lors de l'estimation du pourcentage de sujets à risque, qu'ils ont estimé autour de 25 %. Pour remplir ces conditions, ils devaient inclure 450 sujets par groupe. C'est donc le nombre de sujets nécessaires, pour avoir une précision de 4 % lors de l'évaluation d'une proportion de 25 %.

La notion de représentativité concerne la sélection de la population, mais pas le nombre de sujet nécessaire.

Il s'agissait ici d'une étude transversale pour laquelle aucun test statistique n'a été réalisé. Il n'y a donc pas lieu d'évoquer la significativité statistique.

### Question 12

**Réponses exactes : B – D**

*Commentaires* Dans cette phrase des résultats, les auteurs indiquent qu'un certain nombre de patients, qui étaient considérés à risque de maladie thromboembolique veineuse, présentaient un risque élevé de saignement, contre indiquant la mise en place d'une prophylaxie par anticoagulation. Comment considérer ces patients ?

- En les incluant dans le calcul de la proportion de sujets recevant une prophylaxie adaptée, les auteurs ont commis une erreur. Car bien qu'ils soient à risque, et donc éligibles à une prophylaxie contre la maladie thromboembolique, ces patients ne devaient pas être traités. Le rapport bénéfice risque est jugé défavorables pour eux, en raison d'un risque trop important de saignement.
- En les incluant dans le dénominateur pour le calcul de la proportion (proportion = nombre de patient réellement traités/nombre de patients qui auraient dû être traités), les auteurs sous-entendent que ces patients auraient quand même dû être traités. Il en résulte une sous-estimation de la proportion de patients recevant une prophylaxie adaptée (augmentation du dénominateur = diminution de la proportion).
- Le fait d'être à risque de saignement ne veut pas dire que tous ces patients ont présenté un saignement durant l'étude (effet indésirable médicamenteux).
- La présence d'un haut risque de saignement ne figurait pas dans les critères d'exclusion.

### Question 13

**Réponses exactes : B – E**

*Commentaires* Commentaires : Il s'agit d'une étude transversale descriptive de niveau de preuve faible : niveau 4. Les niveaux de preuves sont définis par la HAS (http://www.has-sante.fr/portail/upload/docs/application/pdf/2013-06/etat_des_lieux_niveau_preuve_gradation.pdf) en fonction du type d'étude.

### Question 14

**Réponses exactes : A – B – E**

*Commentaires*

A, B et C : La loi du 29 décembre 2011 relative au renforcement de la sécurité sanitaire du médicament et des produits de santé impose la transparence des liens entre les industries de santé et les médecins. La déclaration de lien d'intérêt est donc obligatoire. Cette loi impose également la transparence de ces liens : les déclarations sont donc publiques.

D : La puissance d'une étude (*cf.* question 11) est une probabilité, issue de calculs statistiques. Elle ne dépend pas des auteurs, et n'est donc pas impactée par les liens d'intérêts.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

E : Lorsqu'il existe des liens d'intérêts, l'objectivité des auteurs peut être influencée. Attention, ce n'est pas une « obligation » : lien d'intérêt ne veut pas obligatoirement dire « étude faussée ». C'est cependant un risque, qu'il convient d'avoir à l'esprit lors de la lecture critique d'un article scientifique.

## Question 15

**Réponses exactes : A – C – D**

*Commentaires* Pour cette question, il convient de distinguer validité interne et externe d'une étude.

La validité interne concerne la qualité propre de l'étude : absence de biais, qualité méthodologique, qualité de réalisation...

La validité externe concerne la généralisation des résultats de l'étude en question, à la population « générale » (externe à l'étude).

A : La connaissance de cette étude par les médecins, peut avoir influencé (et biaisé) leurs habitudes de prescription. Ce phénomène porte le nom de « effet Hawthorne ». Etant au courant de cette étude et du fait qu'ils étaient observés/jugés, ils ont peut être « mieux » traité les patients à risque. Il y a donc une surestimation des patients bien traités. Cette erreur est propre à cette étude, elle impacte sa qualité. C'est donc de la validité interne.

B : La notion de représentativité renvoie à la généralisation des résultats. Cela influence donc la validité externe.

C et D : Ces deux propositions renvoient à la qualité propre de l'étude : problèmes méthodologiques lors du recueil, erreur de saisie. C'est de la validité interne.

E : La faible présence d'une catégorie de patients n'affecte pas la qualité interne de l'étude. En revanche, lors de la généralisation des résultats, la représentativité de l'étude pourra être questionnée. C'est de la validité externe.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 10 Étude descriptive (sur registre)

Guy Launois

| | |
|---|---|
| **Titre de l'article** | **Survival differences between patients with Hodgkin lymphoma treated inside and outside clinical trials. A study based on the EORTC-Netherlands Cancer Registry linked data with 20 years of follow-up** |
| **Auteurs** | Liu L, Giusti F, Schaapveld M, Aleman B, Lugtenburg P, Meijnders P, Hutchings M, Lemmens V, Bogaerts J, Visser O |
| **Référence** | Br J Haematol. 2017 ; 176(1):65-75. |
| **Lien vers l'article** | http://onlinelibrary.wiley.com/doi/10.1111/bjh.14379/epdf |

**1** **Il s'agit d'une étude**

a. Prospective
b. Observationelle
c. Unicentrique
d. Écologique
e. Conduite en population

**2** **Concernant les procédures réglementaires**

a. Un consentement éclairé a été obtenu auprès de tous les patients
b. La procédure de croisement des fichiers a été anonyme
c. L'analyse des données a été anonymisée
d. Le protocole a été soumis à un ou plusieurs comités scientifiques
e. Le protocole a été soumis à un ou plusieurs comités indépendants des institutions impliquées dans la recherche

**3** **La population de l'étude**

a. Est constituée de tous les patients enregistrés au Registre national des cancers durant les années de l'étude
b. Est constituée de 5 714 personnes
c. Est représentative des patients de moins de 70 ans ayant eu un lymphome comme premier cancer aux Pays-Bas durant les années de l'étude
d. Est constituée des patients ayant eu un lymphome et inclus dans un essai clinique durant les années de l'étude
e. Exclut les patients âgés de plus de 70 ans au moment du diagnostic de leur cancer



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **À propos de l'analyse des données de l'étude**

a. La date de point de l'étude de survie a été fixée au 1er février 2014
b. Des analyses univariées ont été conduites
c. Des analyses multivariées ont été conduites
d. L'analyse statistique n'a porté que sur 5 389 personnes
e. Le critère de jugement était la survie sans récidive

**5** **Concernant l'analyse des données de l'étude**

a. Le Hazard-Ratio a permis de mesurer les ratio de risque instantané de décès entre les patients recrutés dans les essais cliniques et ceux qui ne l'étaient pas
b. Une analyse complémentaire a été conduite pour chaque stade d'extension au diagnostic
c. Le Hazard-Ratio a été utilisé parce que l'analyse a été stratifiée sur le stade d'extension au diagnostic
d. Une analyse complémentaire a été conduite pour chaque groupe d'âge
e. Le Hazard ratio n'aurait pas pu être utilisé si l'analyse n'avait pas été ajustée sur l'âge

**6** **Le tableau I montre que**

a. Les données manquantes sont plus fréquentes dans la population non incluse dans les essais cliniques
b. La population incluse dans les essais cliniques n'est pas plus jeune que celle des patients non inclus
c. Les hôpitaux universitaires recrutent plus dans les essais cliniques que les hôpitaux généraux
d. Le statut socioéconomique est un facteur de confusion potentiel
e. Le recrutement dans les essais cliniques a augmenté significativement au cours du temps

**7** **Concernant les résultats**

a. La grande différence d'effectif entre les 2 groupes (inclus/non inclus) compromet la validité interne de l'étude
b. À la lecture du tableau I, l'âge apparaît comme un facteur de confusion potentiel
c. À la lecture du tableau I, le stade au diagnostic apparaît comme un facteur de confusion potentiel
d. Le pourcentage plus important de données manquantes dans le groupe « non participants » fait du sous type histologique un facteur de confusion
e. Le pourcentage plus important de données manquantes dans le groupe « non participants » oblige à réaliser une stratification sur le sous type histologique



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**8** **La figure 1 montre que**

a. La survie brute globale des patients recrutés dans les essais cliniques est meilleure que celle des non recrutés
b. La survie brute globale des patients jeunes est meilleure que celle des patients âgés
c. La différence de survie entre les 2 groupes (inclus/non inclus) est expliquée en partie par la différence de structure d'âge de ces 2 groupes
d. La population de référence utilisée dans la standardisation était celle des patients non inclus dans les essais cliniques
e. Les patients recrutés dans les essais cliniques sont plus âgés que ceux non recrutés

**9** **La figure 2 montre**

a. La survie ajustée sur l'âge
b. La survie ajustée sur le stade
c. La survie standardisée sur l'âge
d. Montre la survie standardisée sur l'âge en fonction de leur inclusion dans un essai thérapeutique et le stade d'extension de leur cancer
e. Montre la survie stratifiée sur l'âge

**10** **La Figure 3 montre que**

a. La survie brute (non ajustée) globale est légèrement meilleure chez les hommes que chez les femmes
b. La survie globale (non ajustée) est meilleure chez les participants aux essais cliniques chez les hommes comme chez les femmes
c. La survie brute (non ajustée) globale est meilleure dans les stades I que dans les stades II et III
d. Dans la population globale de l'étude, la survie brute (non ajustée) globale n'est pas significativement meilleure chez les patients inclus dans les essais thérapeutiques
e. La survie brute (non ajustée) globale des patients inclus dans les essais thérapeutiques s'est amélioré au cours du temps

**11** **La figure 4**

a. Montre les différences de survie brute (non ajustée) globale entre les patients inclus et non-inclus dans les essais thérapeutiques
b. Montre la différence des taux de survie à 20 ans entre les patients inclus et non-inclus dans les essais thérapeutiques après ajustement sur tous les facteurs de confusion
c. Montre que, après ajustement, le ratio des risques instantanés de décès entre les 2 groupes n'est significatif dans aucune des strates du stade d'extension
d. Montre que, après ajustement, les patients sans traitement ont une survie significativement moins bonne que les patients traités
e. La meilleure survie du groupe inclus dans les essais thérapeutiques constatée dans la strate « sans traitement » suggère la non prise en compte de certains facteurs de confusion résiduels



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**12** **La Figure 5 montre que**

a. La proportion de patients traités par chimio et radiothérapie combinée a augmenté au cours du temps
b. Les essais cliniques les plus récents ont plus porté sur la chimio et radiothérapie combinée que les essais cliniques anciens.
c. L'amélioration de la survie au cours du temps ne concerne que les patients inclus dans les essais thérapeutiques
d. La survie est meilleure chez les patients traités par chimio et radiothérapie combinée
e. La survie est meilleure chez les patients traités par chimio et radiothérapie combinée en dehors des essais thérapeutiques

**13** **Le fait que cette étude utilise des données de registre de cancers**

a. Facilite l'extrapolation des résultats à l'ensemble de la population
b. Conforte la validité externe de l'étude
c. Augmente la puissance de l'étude
d. Conforte la représentativité de la population de l'étude
e. Conforte la comparabilité des groupes

**14** **Concernant la discussion de l'article**

a. La différence de survie globale entre les deux groupes est expliquée en partie par des différences d'âge entre les deux groupes
b. La différence de survie entre les deux groupes est expliquée en partie par un biais de sélection dans l'étude
c. La différence de survie entre les deux groupes est expliquée en partie par les critères d'inclusion dans les essais thérapeutiques
d. La différence de survie entre les deux groupes est expliquée en partie par des différences de stade d'extension de la maladie entre les deux groupes
e. La différence de survie entre les deux groupes est expliquée en partie une augmentation du nombre de patients inclus dans les essais thérapeutiques au cours du temps

**15** **Limites et forces de l'étude**

a. La plus faible précision des données des registres par rapport aux essais cliniques limite l'interprétation des résultats
b. L'utilisation des données du registre a permis d'étudier de manière fiable des données couvrant une longue période de temps.
c. Sans registre, il était impossible de réaliser l'étude de survie
d. La possibilité de croisement nominatif des fichiers a amélioré la fiabilité des résultats
e. L'utilisation des données du registre limite la transférabilité des résultats à d'autres pays



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 1

**Réponses exactes : B – E**

*Commentaires*

A. Faux : l'étude s'appuie sur les bases de données des essais cliniques sur le lymphome de Hodgkin de l'EORTC et sur le registre des cancers des Pays-Bas. L'étude est donc rétrospective.

B. Vrai : il s'agit d'une étude observationnelle (aucune intervention modifiant la prise en charge des patients n'est proposée).

C. Faux : Cette proposition est fausse à double titre : les essais utilisés dans l'étude sont des essais multicentriques et l'identification des cas du registre s'appuie sur les données de l'ensemble des laboratoires d'anatomopathologie des Pays-Bas, ainsi que sur les données de sortie des hôpitaux, les données des laboratoires d'hématologie et des instituts de radiothérapie.

D. Faux : Les données individuelles des patients sont bien utilisées ici.

E. Vrai : cette étude s'appuie sur des données recueillies dans l'ensemble de la population des Pays-Bas pour ce qui concerne le registre des cancers (1er paragraphe, 2e colonne, page 2) et donc par définition est une étude « population-based ».

## Question 2

**Réponses exactes : C – D**

*Commentaires*

A. Faux : un consentement n'a pu être obtenu des patients inclus dans les essais (*cf.* fin page 2-début page 3).

B. Faux : le chaînage a été effectué par l'institut national du cancer, à partir des bases de données nominatives de l'EORTC et du registre des cancers. Les données n'ont été anonymisées qu'à la suite de cette opération avant transfert à l'équipe de l'EORTC en charge de cette analyse.

C. Vrai : *cf.* commentaire précédent.

D. Vrai : L'étude a été validée par le comité consultatif de l'Institut du cancer et par l'IRB de l'EORTC.

E. Faux : Les deux comités en question ne peuvent pas être considérés comme indépendants des institutions impliquées dans la recherche même s'ils sont constitués en partie de personnes extérieurs à ces institutions.

## Question 3

**Réponses exactes : B – C – E**

*Commentaires*

A. Faux : Pour les années antérieures à 1989, les données d'un registre régional ont été utilisées (page 3, 1re colonne, 2e paragraphe).

B. Vrai *cf.* texte.

C. Vrai : L'étude sur la base de données du registre national et sur 3 années antérieures sur la base d'un registre parcellaire. Par ailleurs, les patients ayant un lymphome et inclus dans un essai entre 1986 et 2000 l'ont tous été dans le cadre d'un essai mené par l'EORTC. Seulement 59 patients ont été inclus dans des essais provenant d'un autre groupe entre 2000 et 2004. Ces éléments conduisent tous à penser que la population d'étude est réellement représentative des patients ayant eu un lymphome aux Pays-Bas.

D. Faux : L'étude inclus bien évidemment des patients ayant eu un lymphome, inclus ou non dans un essai clinique durant les années 1986 à 2004. L'inclusion de patients non inclus dans un essai est indispensable pour la comparaison envisagée.

E. Vrai : c'est un critère d'inclusion dans l'étude.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 4**

**Réponses exactes : A – B – C – D**

*Commentaires*

A. Vrai : *cf.* texte (page 3, 1re colonne, 2e paragraphe).

B. Vrai : *cf.* texte page 3, 2e colonne, 3e paragraphe ; on parle aussi d'analyses bivariées.

C. Vrai : *cf.* texte. Ce type d'analyse permet d'estimer l'effet de l'inclusion dans un essai indépendamment des autres facteurs influençant le pronostic des patients.

D. Vrai.

E. Faux : le critère de jugement était la survie globale.

**Question 5**

**Réponses exactes : A – B – D**

*Commentaires*

A. Vrai à condition de préciser l'item sinon me paraît ambigu. Le Hazard Ratio (ou Risque Relatif) correspond bien à l'estimation moyenne sur l'ensemble du suivi du ratio du risque instantané de décès des patients recrutés dans un essai et des patients non recrutés.

B. Vrai *cf.* résultats.

C. Faux : le Hazard ratio est calculable que l'analyse soit stratifiée ou non sur un facteur tiers.

D. Vrai *cf.* figure 3.

E. Faux : Le Hazard ratio peut être estimé dans le cadre d'une analyse univariée et pas uniquement multivariée.

**Question 6**

**Réponses exactes : A – C**

*Commentaires* Commentaires

Vrai *cf.* texte.

Faux, *cf.* tableau 1, les patients inclus dans les essais sont plus jeunes.

Vrai *cf.* tableau 1 49,5 % vs 39,7 %, $p < 10\text{-}4$.

Faux *cf.* tableau 1. Pour être un facteur de confusion il est nécessaire que la variable considérée soit associée avec l'événement d'intérêt (ici la survie) ET avec la variable d'exposition étudiée (ici la participation à un essai).

Faux : Le pourcentage de patients inclus ne diffère pas significativement d'une période de diagnostic à l'autre $p = 0{,}297$.

**Question 7**

**Réponses exactes : B – C**

*Commentaires*

A. Faux : la validité interne dépend essentiellement du contrôle des biais (sélection, mesure, confusion) même si elle dépend aussi du contrôle de l'erreur aléatoire et donc de la taille de l'échantillon. La différence d'effectif entre les groupes est certes importante, mais le nombre de patients inclus dans les essais est suffisant pour ne pas menacer la précision des estimations.

B. Vrai, *cf.* tableau 1, l'âge est lié à l'inclusion dans un essai (les plus jeunes sont davantage inclus) et avec la survie. Autre indice les auteurs ont ajusté leurs analyses sur ce facteur.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

C. Le raisonnement est le même pour le stade. Cependant une hétérogénéité significative est mise en évidence dans la figure 3, laissant suspecter également une association différente entre inclusion dans un essai et survie en fonction du stade (interaction).

D. Faux *cf.* tableau 1 Ceci peut contribuer à expliquer partiellement l'association entre type histologique et inclusion dans un essai (différence de pourcentage d'inclusion en fonction du type histologique), mais le différentiel sur les données manquantes ne suffit pas à elle seule.

E. Faux : Le pourcentage plus important de données manquantes chez les patients non inclus n'oblige pas à une stratification. La stratification est rendue nécessaire si l'association entre inclusion dans un essai et survie varie en fonction du type histologique (mise en évidence d'une interaction ou d'une hétérogénéité).

## Question 8

**Réponses exactes : A – C – D**

*Commentaires*

A. Vrai : d'après la figure 1 la survie brute globale (crude OS) des patients inclus dans les essais (survie brute à 10 ans estimée à 85 % d'après la figure et le texte) apparaît effectivement meilleure que celle des patients non inclus (77 %). Cette information est confirmée par l'estimation du RR brut (OR = 0,75 ; 95 % -CI = 0,68-0,83). Dans les études en population (population based) reposant sur des données de registre, on utilise le terme de survie « brute » (crude) pour différentier cette estimation de la survie spécifiquement liée au cancer. L'adjectif « brut » associé au risque relatif signifie en revanche qu'il s'agit du RR non ajusté obtenu par une analyse bivariée et non multivariée.

B. Faux : la figure 1 ne permet pas de montrer cela.

C. Vrai : la notion de survie standardisée sur l'âge correspond à une estimation de la survie dans laquelle on estime la survie par âge et on l'applique à une population externe de structure d'âge connue ou à la population à laquelle on veut comparer la population d'étude. Cette façon de faire permet de s'affranchir de différences de structure d'âge entre les populations comparées. Dans la figure 1, les différences entre les survies des patients inclus et non inclus sont largement diminuées par la standardisation, la survie à 10 ans passant de 81 % à 77 % soit une réduction de moitié de la différence observée en survie brute (*cf.* A).

D. Vrai : on peut déduire cette information du fait que la courbe de la survie brute etde la survie standardisée dans le groupe non inclus sont superposées.

E. Faux : La figure 1 ne donne pas d'indication directe sur les différences d'âge. Pour cela il est nécessaire de se référer au tableau 1.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 9

**Réponses exactes : C – D**

*Commentaires*

A. Faux : La standardisation est une méthode différente de l'ajustement. Un ajustement nécessite l'application d'un modèle multivarié (par exemple modèle de Cox) incluant à la fois la variable inclusion dans un essai et l'âge. Ce modèle sous certaines hypothèses peut permettre d'estimer la survie dans le groupe des patients inclus *versus* celle des non inclus à âge égal.

B. Faux : Les courbes de survie sont estimées stade par stade ; elles ne sont pas ajustées sur le stade.

C. Vrai il s'agit bien de courbes de survie standardisées sur l'âge présentées pour les patients inclus et non inclus et pour chaque stade.

D. Vrai C'est exactement ce que représentent les composantes de la figure 2.

E. Faux : La figure 2 ne donne pas de courbes par tranche d'âge.

## Question 10

**Réponse exacte : B**

*Commentaires*

A. Faux : La figure 3 présente les risques relatifs des patients inclus vs non inclus dans les essais chez les hommes et chez les femmes, mais ne compare pas la survie des hommes et des femmes.

B. Vrai : Les intervalles de confiance des risques relatifs chez les hommes et les femmes n'incluent pas 1, l'affirmation est donc vraie.

C. Faux *cf.* A.

D. Faux : l'estimation globale fait apparaître un RR = 0,75 avec un intervalle de confiance entre 0,68 et 0,83, ne comprenant pas 1 ; la survie des patients inclus dans les essais thérapeutiques est donc significativement améliorée par rapport à celle des patients non recrutés pour un essai durant la période d'étude.

E. Faux la figure 3 ne permet pas d'affirmer cela. Le test d'hétérogénéité est de plus non significatif, indiquant qu'on ne peut pas dire que l'effet sur la survie globale d'être inclus dans un essai s'est modifié au cours du temps.

## Question 11

**Réponses exactes : C – E**

*Commentaires*

A. Faux : La légende indique qu'il s'agit d'estimations issues d'une analyse multivariée.

B. Faux : la figure représente des risques relatifs, donc des ratio moyens de risque de décès chez les patients inclus vs non inclus dans des essais ; il ne s'agit pas de différence de taux de survie.

C. Vrai. Aucun des intervalles de confiance mentionnés pour chaque stade n'exclut la valeur 1, et donc aucune relation significative n'a été montrée.

D. Faux : chez les patients non traités, le risque relatif correspondant à la comparaison de la survie/mortalité des patients inclus et non inclus dans un essai ; ce risque relatif bien que significatif ne permet pas d'affirmer que les patients sans traitement ont une survie significativement moins bonne que les patients traités.

E. Vrai il est effectivement probable que la moins bonne survie des patients inclus vs non inclus dans le groupe non traité s'explique par des facteurs de confusion non pris en compte dans cette étude d'autant que les essais en cancérologie sont exceptionnellement contre placebo.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 12

**Réponses exactes : A – B**

*Commentaires*

A. Vrai : C'est exactement ce que montre la figure 5B.

B. Vrai : le fait que la proportion de patients traités en 1re intention par chimioradiothérapie augmente au cours du temps dans le cadre des essais suggère que cette affirmation est vraie ; cette affirmation est corroborée par l'indication page 8, 2e colonne que notamment les essais H6 à H9 incluent différentes stratégies combinant radio et chimiothérapie.

C. Faux. D'après la figure 5A l'amélioration de la survie concerne aussi bien les patients inclus dans les essais que les patients non inclus.

D. Faux : la figure 5 ne donne aucune indication sur cela.

E. Faux : même raison qu'item D.

### Question 13

**Réponses exactes : A – B – D**

*Commentaires*

A. Vrai : L'utilisation d'un chaînage avec le registre des cancers limite certains biais de sélection

B. Vrai : l'utilisation de données de population facilite l'extrapolation des résultats à l'ensemble de la population cible.

C. Faux. La puissance de l'étude ne dépend pas directement de l'utilisation d'une base de données spécifiques mais de sa taille.

D. Vrai : idem B.

E. Faux : le fait d'utiliser des données de population ne modifie pas le fait que les patients inclus diffèrent des patients non inclus dans les essais, ne serait-ce que du fait des critères d'éligibilité de ces derniers.

### Question 14

**Réponses exactes : A – C – D**

*Commentaires*

A. Vrai : La standardisation sur l'âge contribue à diminuer de façon importance la différence entre la survie des patients inclus et celle des patients non inclus dans les essais.

B. Faux : l'étude intègre l'ensemble des patients atteints de cancer en Hollande. Seule la période la plus ancienne ne concerne qu'une partie du pays, mais l'impact est sans doute négligeable.

C. Vrai. La plupart des essais thérapeutiques limite l'inclusion de sujets avec des comorbidités par exemple.

D. Vrai : *cf.* page 6, 1re colonne ; les essais incluent plus de patients à une phase précoce de leur pathologie.

E. Faux : cette affirmation est fausse ; si le % de patient inclus dans les essais augmente, cela est indépendant de la survie dans chacun des groupes et donc de la différence entre les deux. En pratique, l'augmentation des inclusions peut résulter d'un élargissement des critères d'inclusion et dans ce cas la différence entre les caractéristiques initiales entre les deux groupes va être réduite, contribuant à réduire la différence ; soit l'augmentation des inclusions se fait en optimisant le taux d'acceptation des patients sans élargir forcément les critères d'éligibilité, et cela n'a alors aucun impact sur la différence de survie.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

■ **Question 15**

**Réponses exactes : A – B – D**

*Commentaires*

A. Vrai : par exemple il n'a pas été possible d'ajuster sur les comorbidités par exemple et de ce fait l'interprétation ne peut tenir compte de cet élément.

B. Vrai : l'existence des données de registre sur la période 1986-2004 a permis de couvrir cette période de manière faible, du fait du contrôle permanent de l'exhaustivité de l'identification des cas de cancer dans le registre. Il n'aurait pas été possible de réaliser une étude fiable sans cela.

C. Faux. Il est tout à fait possible de mener une étude de survie sans registre. Le registre facilite permet d'identifier de façon fiable les patients atteints de cancer même non inclus dans un essai, mais ce n'est pas la seule possibilité d'identifier le statut vital des patients.

D. Vrai : la possibilité de croiser les bases de données de façon nominative réduit les risques d'erreur par rapport à un croisement sur des données non nominatives reposant par exemple sur la date de naissance, le sexe et le code postal de résidence.

E. Faux : ce sont les caractéristiques de la population qui peuvent limiter la transférabilité, mais pas le fait que l'on utilise des données de registre. La seule chose qui pourrait limiter la transférabilité des résultats à d'autres pays, est que les patients atteints d'un lymphome de Hodgkin au Pays-Bas soient très différents, pour des raisons génétiques par exemple, des patients atteints d'un lymphome de Hodgkin d'un autre pays. Ceci est hautement improbable ici.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 11 Étude diagnostique

Laurent Boyer, Noémie Resseguier

| | |
|---|---|
| **Titre de l'article** | **Computed Tomographic Colonography (Virtual Colonoscopy) : A Multicenter Comparison With Standard Colonoscopy for Detection of Colorectal Neoplasia** |
| **Auteurs** | Cotton P., Durkalski V., Pineau B., Palesch, Y., Mauldin P., Hoffman B., Vining D., Small W., Affronti, J., Re D., Kopecky K., Ackerman S., Burdick J., Brewington C., Turner M., Zfass A., Wright A., Lyer R., Lynch P., Sivak M., Butler H |
| **Référence** | JAMA. 2004 ; 291(14):1713-1719. |
| **Lien vers l'article** | https://jamanetwork.com/journals/jama/fullarticle/198538 |

**1** **Il s'agit d'une étude**

a. Randomisée
b. De supériorité
c. En cross-over
d. De type méta-analyse
e. Diagnostique

**2** **L'objectif principal de cette étude était d'évaluer**

a. La fiabilité de la coloscopie virtuelle sur un large effectif réparti dans plusieurs centres
b. La validité de la coloscopie virtuelle sur un large effectif réparti dans plusieurs centres
c. De déterminer la prévalence du cancer colorectal sur un large effectif réparti dans plusieurs centres
d. De déterminer l'incidence du cancer colorectal sur un large effectif réparti dans plusieurs centres
e. De réduire la mortalité du cancer colorectal en le diagnostiquant précocement sur un large effectif réparti dans plusieurs centres

**3** **Les auteurs indiquaient dans l'introduction qu'il existait des variations importantes entre les différents examinateurs dans l'interprétation de la coloscopie virtuelle. Les auteurs auraient pu proposer une analyse complémentaire et évaluer**

a. La fiabilité de la coloscopie virtuelle
b. La reproductibilité de la coloscopie virtuelle
c. La vraisemblance de la coloscopie virtuelle
d. La validité externe de la coloscopie virtuelle
e. La concordance inter-juge de la coloscopie virtuelle



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **À propos de l'un des critères de jugement principal, la sensibilité**

a. Est un indicateur extrinsèque de validité de la coloscopie virtuelle
b. Est la probabilité conditionnelle que la coloscopie virtuelle soit positive si le sujet est atteint de la maladie
c. Mesure la capacité de la coloscopie virtuelle à bien classer les sujets sains
d. Permet d'exclure la maladie lorsqu'elle est égale à 100 % et que la coloscopie virtuelle est négative
e. Varie avec la prévalence de la maladie

**5** **À propos de l'un des critères de jugement principal, la spécificité**

a. Est un indicateur intrinsèque de validité de la coloscopie virtuelle
b. Est la probabilité conditionnelle que la coloscopie virtuelle soit positive si le sujet est atteint de la maladie
c. Mesure la capacité de la coloscopie virtuelle à bien classer les sujets sains
d. Permet de confirmer la maladie lorsqu'elle est égale à 100 % et que la coloscopie virtuelle est négative
e. Varie avec la prévalence de la maladie

**6** **Les critères de jugement secondaires de cette étude comprenaient**

a. Des indicateurs intrinsèques de validité de la coloscopie virtuelle
b. Des indicateurs extrinsèques de validité de la coloscopie virtuelle
c. Des indicateurs de concordance entre la coloscopie virtuelle et la coloscopie standard
d. Les préférences des professionnels entre la coloscopie virtuelle et la coloscopie standard
e. Le taux de détection correcte des lésions

**7** **L'examen de référence (ou gold-standard) choisi dans cette étude pour le diagnostic des lésions était**

a. La coloscopie classique
b. La coloscopie classique associée à la coloscopie virtuelle
c. L'examen anatomo-pathologique
d. Une combinaison des résultats initiaux de la coloscopie classique et des résultats d'éventuels examens complémentaires réalisés après que l'endoscopiste ait eu connaissance des résultats de la coloscopie virtuelle.
e. Il n'y avait pas d'examen de référence



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**8** **Les auteurs réalisaient une levée d'insu des résultats de la coloscopie virtuelle. Les endoscopistes n'avaient pas connaissance des résultats de la coloscopie virtuelle au moment de l'insertion du coloscope. Une fois que le segment colique avait été examiné et que les constatations de l'examen avaient été notées, l'endoscopiste prenait connaissance des résultats de la coloscopie virtuelle, ce qui permettait un réexamen immédiat en cas de désaccord. Cette levée d'insu est justifiée**

a. Parce que la coloscopie standard est imparfaite
b. Parce que la coloscopie virtuelle est imparfaite
c. Parce que la levée d'insu permet d'éviter de sous-estimer les performances diagnostiques de la coloscopie virtuelle
d. Parce que la levée d'insu tente d'éviter une qualification erronée en « faux positifs » des lésions omises par une 1re coloscopie standard
e. Éthiquement

**9** **Concernant le tableau 2 présentant les résultats portant sur l'unité expérimentale « par lésion », les auteurs ne présentaient pas les résultats relatifs à la spécificité**

a. Il est possible de calculer la spécificité avec l'unité expérimentale « par lésion » mais celle-ci n'est pas présentée dans le tableau
b. La spécificité étant égale à 1- sensibilité, il est acceptable que ce résultat ne soit pas présenté
c. Parce qu'il n'est pas possible de calculer la spécificité avec l'unité expérimentale « par lésion » car il n'est pas possible de dénombrer les lésions absentes
d. L'unité expérimentale « par patient », contrairement à l'unité « par lésion », permet de définir les patients « vrais négatifs », et possède donc une double expression possible en sensibilité et spécificité
e. L'unité expérimentale « par lésion », contrairement à l'unité « par patient », permet de définir les patients « vrais négatifs », et possède donc une double expression possible en sensibilité et spécificité

**10** **L'unité expérimentale « par patient » est soumise à un risque de**

a. Biais de sélection
b. Biais de publication
c. Biais d'attrition
d. Biais de confusion
e. Cette unité expérimentation n'est soumise à aucun biais particulier

**11** **Dans les tableaux 1 et 2, l'estimation des sensibilités est accompagnée d'intervalles de confiance à 95 %. L'intervalle de confiance renseigne sur la valeur de la sensibilité de la coloscopie virtuelle dans**

a. L'échantillon étudié
b. La population d'analyse
c. La population source
d. La population cible
e. La population type



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**12** **L'une des forces de cette étude était sa réalisation multicentrique (9 centres) par rapport aux études précédentes qui étaient majoritairement monocentriques. Il s'agit d'une force de cette étude car**

a. Cela a permis l'augmentation de la capacité de recrutement
b. Cela a amélioré la possibilité de généralisation des résultats, en prenant en compte la variabilité intercentre
c. Cela a amélioré la validité externe de l'étude
d. Cela a amélioré la validité interne de l'étude
e. Cela a permis d'éviter tout biais de sélection

**13** **L'analyse de préférence auprès des patients entre la coloscopie virtuelle et la coloscopie standard renseigne**

a. La validité intrinsèque de la coloscopie virtuelle
b. La validité extrinsèque de la coloscopie virtuelle
c. L'étude des propriétés discriminantes de la coloscopie virtuelle
d. La fiabilité de la coloscopie virtuelle entre les patients et les cliniciens
e. L'acceptabilité de la coloscopie virtuelle

**14** **Dans cette analyse de préférence**

a. Tous les patients ont répondu à l'enquête des préférences
b. 518 patients sur 600 ont répondu à l'enquête des préférences
c. 46 % des patients ont répondu à l'enquête des préférences
d. 41 % des patients ont répondu à l'enquête des préférences
e. 13 % des patients ont répondu à l'enquête des préférences

**15** **La proportion de non-répondants était de 13,7 % dans l'analyse des préférences. Ce taux de non-participation peut avoir pour conséquence**

a. Un biais lorsque les non-répondants ont un comportement différent des répondants
b. Une baisse de la qualité de l'inférence
c. Que l'estimateur obtenu admet une variance d'échantillonnage plus grande (il est donc moins précis) puisque la taille de l'échantillon exploitable est plus faible que celle prévue initialement
d. Les auteurs auraient pu comparer les caractéristiques des répondants et des non-répondants afin de s'assurer qu'ils ne différaient pas sur des critères sociodémographiques et cliniques.
e. Il n'y a pas de biais de non-réponse dans cette étude



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 1**

**Réponse exacte : E**

*Commentaires* Il s'agit d'une étude diagnostique (réponse E vraie et réponse D fausse). La méthodologie de l'étude est décrite dans la partie « study design » et précise qu'il s'agit d'une étude non randomisée (réponse A fausse), selon un design de non-infériorité (réponse B fausse). Il n'y a pas de cross over dans cette étude (réponse C fausse).

**Question 2**

**Réponse exacte : B**

*Commentaires* Il s'agit d'une étude évaluant la validité (accuracy, définie comme la capacité de différencier au sein de la population cible les personnes probablement atteintes de la maladie de celles qui sont probablement indemnes) de la coloscopie virtuelle sur un large effectif réparti dans plusieurs centres. Les réponses A, C, D et E ne correspondent pas à l'objectif de l'étude « To assess the accuracy of CTC in a large number of participants across multiple centers ».

**Question 3**

**Réponses exactes : A – B – E**

*Commentaires* « Fiabilité » et « reproductibilité » sont synonymes. Un test reproductible est un test qui donne le même résultat sur un sujet, quel que soit l'opérateur (reproductibilité interopérateur) et quel que soit le moment où on l'applique (reproductibilité temporelle/intra opérateur). Plus un examen est reproductible, plus il est fiable. Les réponses A, B et E sont donc justes puisque l'introduction insistait sur les variations de résultats entre « lecteur », il aurait donc été intéressant d'étudier la fiabilité, reproductibilité ou concordance des résultats entre les lecteurs (ou juges). Le coefficient kappa est une mesure permettant d'évaluer le taux de concordance inter-observateurs pour deux variables qualitatives appariées, alors que la méthode de Bland et Altman est adaptée pour les variables quantitatives. Les réponses C et D relèvent de la validité du test.

**Question 4**

**Réponses exactes : B – D**

*Commentaires* La sensibilité est un indice intrinsèque de validité de la coloscopie virtuelle (réponse A fausse). La sensibilité est la probabilité conditionnelle que la coloscopie virtuelle soit positive si le sujet est atteint de la maladie (Se = P(T+/M+), réponse B vraie). Elle mesure donc la capacité de la coloscopie virtuelle à bien classer les malades (réponse C fausse). Lorsque la sensibilité est égale à 100 %, il n'existe aucun faux négatif (réponse D vraie). Elle ne varie pas avec la prévalence de la maladie (réponse E fausse). Seuls les indicateurs extrinsèques varient avec la prévalence.

**Question 5**

**Réponses exactes : A – C**

*Commentaires* La spécificité est un indice intrinsèque de la coloscopie virtuelle (réponse A vraie). La spécificité est la probabilité conditionnelle que l'examen soit négatif si le sujet est indemne de la maladie (Sp = P(T-/M-), réponse B fausse). Elle mesure donc la capacité de la coloscopie virtuelle à bien classer les non-malades (réponse C vraie). Lorsque la spécificité est égale à 100 %, il n'existe aucun faux positif. Si l'examen est positif, les sujets sont dont atteints de la maladie étudiée (réponse D fausse). Elle ne varie pas avec la prévalence de la maladie (réponse E fausse).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 6

**Réponses exactes : B – E**

*Commentaires* La sensibilité et la spécificité sont des indices intrinsèques mais ils correspondaient au critère de jugement principal (réponse A fausse). Les valeurs prédictives positives et négatives sont évaluées dans les objectifs secondaires et correspondent à des indicateurs extrinsèques (réponse B vraie). La concordance n'était pas mesurée dans cette étude (réponse C fausse). Les préférences étaient étudiées mais il s'agissait de celles des patients et non des professionnels (réponse D fausse). La réponse E était vraie (« secondary outcome measures included the correct detection rate for lesions… »).

### Question 7

**Réponse exacte : D**

*Commentaires* Il y a toujours un gold-standard ou examen de référence dans une étude diagnostique auquel est confronté le nouveau test afin de déterminer ses performances (réponse E fausse). Dans cette étude, il s'agissait d'une combinaison des résultats initiaux de la coloscopie classique et des résultats d'éventuels examens complémentaires réalisés après que l'endoscopiste ait eu connaissance des résultats de la coloscopie virtuelle (réponse D vraie).

### Question 8

**Réponses exactes : A – C – D – E**

*Commentaires* La levée de l'insu est justifiée car la coloscopie standard est imparfaite et ne permet pas dans 100 % des cas un diagnostic de certitude (réponse A vraie et réponse B fausse). De plus la levée d'insu satisfait à des principes éthiques car elle incite à vérifier au moyen d'une deuxième coloscopie l'existence d'une lésion identifiée par la coloscopie virtuelle et non retrouvée par la première coloscopie standard (réponse E vraie).

Lorsqu'un réexamen par coloscopie standard était pratiqué :

- la lésion objectivée par la coloscopie virtuelle était alors qualifiée de « faux positif » si elle n'a pas été retrouvée lors du réexamen par la coloscopie standard. Dans cette situation, les études avec et sans levée d'insu qualifient de la même façon ce résultat de la coloscopie virtuelle ;
- la lésion objectivée par la coloscopie virtuelle était alors qualifiée de « vrai positif » si le réexamen par coloscopie standard confirmait l'omission du premier examen : le gain est alors double pour la coloscopie virtuelle, puisque ce « faux positif » potentiel est in fine qualifié de « vrai positif », augmentant simultanément la sensibilité (augmentation du numérateur) et la spécificité de la coloscopie virtuelle (diminution du dénominateur), ce en comparaison aux études sans levée d'insu (réponses C et D vraies).

### Question 9

**Réponses exactes : C – D**

*Commentaires* Ce tableau ne présente pas la spécificité car il n'est pas possible de la calculer ici (réponse A fausse). En effet, l'unité « per lésion » décrit les capacités de détection par coloscopie virtuelle de chaque lésion colique cible appartenant à une catégorie de taille prédéfinie. Cette unité a pour principal inconvénient de ne pas permettre d'expression de spécificité d'exploration colique : il n'est en effet pas possible de « dénombrer » les lésions absentes (réponse C vraie). Au contraire, l'unité « per patient » permet de définir des patients « vrais négatifs » et possède donc une double expression en sensibilité et spécificité (réponse D vraie et réponse E fausse). Par ailleurs, la spécificité n'est pas égale à 1- sensibilité (réponse B fausse).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 10

**Réponse exacte : A**

*Commentaires* Tout patient pour lequel la coloscopie virtuelle a identifié au moins une lésion parmi celles présentes est qualifié de « vrai positif par patient ». Ainsi, un patient pour lequel la CV identifie un polype de 7 mm, mais n'identifierait pas un polype de 10 mm, est considéré comme « vrai positif per patient » de détection des lésions « ≥ 6 mm ». En unité « per lésion », cet examen de CV aurait abouti à un « vrai positif » (polype de 7 mm) et à un « faux négatif » (polype de 10 mm). L'unité « per patient » représente ainsi potentiellement une expression plus favorable des performances de la CV. La proportion de patients présentant des lésions coliques multiples peut donc influencer la performance diagnostique de la coloscopie virtuelle. L'unité expérimentale « par patient » peut être soumise à un risque de sélection (réponse A vraie) qui serait lié au profil clinique des sujets de l'échantillon différent de celui de la population cible. Les autres biais ne sont pas directement liés au choix de l'unité expérimentale « par patient ».

### Question 11

**Réponse exacte : C**

*Commentaires* L'intervalle de confiance traduit la précision de l'estimation d'un paramètre. Le but de l'estimation est de connaître à partir d'un échantillon la valeur d'un paramètre dans une population, ici la sensibilité de la coloscopie virtuelle. Cependant, l'estimation ponctuelle de ce paramètre sur l'échantillon peut être assez éloignée de la valeur dans la population, du fait des fluctuations aléatoires d'échantillonnage (c'est-à-dire du fait du hasard). L'intervalle de confiance permet de prendre en compte cette incertitude aléatoire dans la présentation des résultats. L'intervalle de confiance à 95 % représente la fourchette de valeurs à l'intérieur de laquelle nous sommes certains à 95 % de trouver la vraie valeur recherchée dans la population source et potentiellement de la cible, si on est sûr que la source soit représentative de la cible. Dans cette étude, l'extrapolation à la population source est la seule à retenir (réponse C vraie) car nous n'avons pas assez d'éléments pour garantir que la population source soit représentative de la population cible.

### Question 12

**Réponses exactes : A – B – C**

*Commentaires* Une étude multicentrique est réalisée simultanément dans plusieurs centres. Les avantages sont l'augmentation de la capacité de recrutement (lorsque le recrutement par un seul centre entraînerait une durée excessive de recrutement) et la possibilité de généralisation des résultats (variabilité inter-centre) (réponses A et B vraies). Elle améliore donc la validité externe de l'étude, c'est-à-dire la généralisation des résultats obtenus (réponse C vraie). Elle a peu d'influence sur la validité interne de l'étude (dépendant du contrôle des différents biais lors de la planification et la conduite de l'étude, réponse D fausse) et ne permet pas d'éviter tout biais de sélection (réponse E fausse).

### Question 13

**Réponse exacte : E**

*Commentaires* L'étude des préférences renseigne l'acceptabilité de la coloscopie virtuelle, c'est-à-dire son intérêt pour un patient ou pour une population (réponse E vraie). La mise en évidence des freins et moteurs à la faisabilité et à l'acceptabilité d'un test est important pour sa diffusion.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

■ **Question 14**

**Réponse exacte : B**

*Commentaires* Tous les patients n'ont pas répondu à cette enquête (réponse A fausse). Le taux de participation était de 86,3 % (518 patients sur 600 qui ont eu les 2 procédures), soit 13,7 % de non-répondants (réponse B vraie).

■ **Question 15**

**Réponses exactes : A – B – C – D**

*Commentaires* Les estimateurs calculés sont biaisés car en l'absence de traitement des non-réponses, on fait implicitement l'hypothèse que le comportement des non-répondants est identique à celui des répondants, ce qui est rarement exact (les non-répondants ont souvent des profils spécifiques qui sont susceptibles d'affecter leur taux de participation) (réponse A vraie) et la précision des estimateurs est plus faible puisque la taille de l'échantillon exploitable est plus petite (réponse B et C vraies). Il est habituel de s'assurer que les caractéristiques socio-démographiques et cliniques disponibles ne diffèrent pas entre les répondeurs et non répondeurs, ce qui est un argument en faveur de l'absence d'un biais de non-réponse (réponse D vraie). On ne peut cependant pas être certain de l'absence de biais de non-réponse dans cette étude même si le taux de participation reste élevé.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 12 Étude diagnostique

Noémie Resseguier, Laurent Boyer

| | |
|---|---|
| **Titre de l'article** | **Accuracy and interobserver agreement between MR-non-expert radiologists and MR-experts in reading MRI for suspected appendicits** |
| **Auteurs** | Leeuwenburgh MM, Wiarda BM, Jensch S, van Es HW, Stockmann HB, Gratama JW, Cobben LP, Bossuyt PM, Boermeester MA, Stoker J, OPTIMPAP Study group |
| **Référence** | Eur J Radiol 2014 ; 83(1):103-110 |
| **Lien vers l'article** | http://www.ejradiology.com/article/S0720-048X(13)00521-4/fulltext |

**1** **Le(s) argument(s) suivant(s) est(sont) avancé(s) par les auteurs pour justifier leur étude**

a. L'intérêt du scanner (CT) par l'absence de néphropathie induite par les produits de contraste iodés
b. L'intérêt de l'imagerie par résonance magnétique (MRI) par l'absence de rayonnement ionisant
c. L'absence d'étude visant à évaluer les performances diagnostiques de l'imagerie par résonance magnétique (MRI) pour le diagnostic d'appendicite
d. L'importance de l'incidence de l'appendicite chez les femmes enceintes et chez les enfants
e. La faible connaissance des propriétés diagnostiques de l'imagerie par résonance magnétique (MRI) lorsqu'elle est interprétée par des radiologues non experts de cette technique

**2** **La (les) proposition(s) suivantes correspond(ent) à l'objectif principal de cette étude**

a. L'évaluation des performances diagnostiques de l'imagerie par résonance magnétique (MRI) par rapport au scanner (CT) pour le diagnostic d'appendicite et d'autres pathologies urgentes
b. La comparaison des performances diagnostiques de l'imagerie par résonance magnétique (MRI) pour le diagnostic d'appendicite et d'autres pathologies urgentes, interprétée par des radiologues experts *versus* interprétée par des radiologues non experts
c. L'évaluation de l'accord inter-juges entre radiologues non experts et radiologues experts lors de l'interprétation de l'imagerie par résonance magnétique (MRI) pour le diagnostic d'appendicite et d'autres pathologies urgentes
d. L'évaluation de l'accord inter-juges chez les non experts lors de l'interprétation du scanner (CT) pour le diagnostic d'appendicite et d'autres pathologies urgentes
e. L'évaluation des performances diagnostiques du scanner (CT) et de l'imagerie par résonance magnétique (MRI) par rapport au gold standard pour le diagnostic d'appendicite et d'autres pathologies urgentes



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**3** **Au total, chaque imagerie par résonance magnétique (MRI) est évaluée de façon indépendante**

a. Trois fois, une fois par un radiologue non expert et deux fois par des radiologues experts différents
b. Trois fois, deux fois par un même radiologue non-expert et une fois par un radiologue expert
c. Huit fois par des radiologues non experts différents
d. Trois fois, deux fois par des radiologues non-experts différents et une fois par deux radiologues experts
e. Quatre fois, deux fois par des radiologues non-experts différents et deux fois par des radiologues experts différents

**4** **Dans le cadre de cette étude, les radiologues non-experts ont bénéficié au préalable d'une formation concernant l'interprétation de l'imagerie par résonance magnétique (MRI) réalisée pour suspicion d'appendicite, ce qui a permis**

a. De limiter le biais de sélection
b. De limiter le biais de mesure
c. De limiter le biais de confusion
d. De diminuer la variabilité inter-individuelle des patients
e. De diminuer la variabilité inter-juge « non expert »

**5** **Les patients inclus et analysés dans cette étude bénéficient du(des) examen(s) suivant(s)**

a. D'une radiographie de l'abdomen sans préparation systématiquement
b. D'une imagerie par résonance magnétique (MRI) systématiquement
c. D'un scanner ou d'une imagerie par résonance magnétique (MRI) en fonction de leur groupe de randomisation
d. D'un scanner systématiquement
e. D'aucun examen si leur tableau clinique est très fortement évocateur d'une appendicite

**6** **Dans cette étude, l'examen défini comme référence ou gold standard se base sur le(s) élément(s) suivant(s)**

a. Les données cliniques
b. Les données d'imagerie par résonance magnétique (MRI)
c. Les données chirurgicales
d. Les données anatomopathologiques
e. Le suivi à trois mois

**7** **Au vu des éléments décrits par les auteurs, le calcul du nombre de sujets nécessaire pour cette étude a fait intervenir**

a. Le risque alpha
b. Le risque bêta
c. La valeur de la sensibilité attendue de l'imagerie par résonance magnétique (MRI) pour le diagnostic d'appendicite aiguë
d. La différence de sensibilité attendue entre les radiologues non-experts et les radiologues experts
e. La largeur de l'intervalle de confiance à 90 % souhaitée



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**8** **Les auteurs précisent les horaires et les jours d'inclusion possibles pour cette étude, car ils ont pu**

a. Être à l'origine d'un biais de confusion
b. Être à l'origine d'un biais de sélection
c. Limiter la validité externe de l'étude
d. Limiter la validité interne de l'étude
e. Limiter la significativité statistique des résultats obtenus

**9** **À propos du diagnostic d'appendicite aiguë dans cette étude, et au vu des données de l'imagerie par résonance magnétique (MRI), il est correct de dire que**

a. 89 % des cas d'appendicite aiguë sont identifiés par les radiologues non experts
b. 3 % des cas d'appendicite aiguë ne sont pas identifiés par les radiologues experts
c. Pour les indicateurs de sensibilité et de spécificité, les performances des radiologues experts peuvent être considérées comme similaires à celles des radiologues non-experts
d. Si un radiologue expert considère qu'un patient présente une appendicite aiguë, la probabilité que ce diagnostic soit exact est de 96 %
e. La fréquence de l'appendicite aiguë dans l'étude est de 52 %

**10** **À propos du diagnostic d'une pathologie urgente dans cette étude, et au vu des données de l'imagerie par résonance magnétique (MRI), il est correct de dire que**

a. Pour les indicateurs de sensibilité et de spécificité, les performances des radiologues experts peuvent être considérées comme similaires à celles des radiologues non-experts
b. L'accord entre radiologues experts et radiologues non experts peut être qualifié d'excellent
c. Il n'y a aucun faux négatif après interprétation par les radiologues experts pour le diagnostic de diverticulite aiguë
d. Si les auteurs avaient calculé le rapport de vraisemblance négatif pour le diagnostic d'occlusion intestinale après interprétation par les radiologues experts, il aurait été estimé à 0, ce qui correspond à un apport diagnostique très fort
e. L'intervalle de confiance à 95 % de la sensibilité pour le diagnostic des troubles des voies urinaires (interprétation par les radiologues experts) est asymétrique, ce qui est une erreur de la part des auteurs

**11** **Parmi les éléments suivants, quel(s) est(sont) celui(ceux) qui auraient pu également être utilisé(s) afin d'évaluer les performances diagnostiques de l'imagerie par résonance magnétique (MRI) lorsqu'elle est interprétée par des radiologues non-experts, pour le diagnostic d'appendicite aiguë ?**

a. Le rapport de vraisemblance positif
b. L'indice de Youden
c. L'aire sous la courbe ROC
d. Le graphique de Bland et Altman
e. La courbe ROC



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**12** **Dans cette étude, il est correct de dire que**

a. Le nombre de patients analysés était de 230
b. Le nombre de patients avec une occlusion intestinale comme diagnostic final était de trois
c. Le nombre de patients non inclus car ils se sont présentés hors des horaires définis pour l'étude, était de 330
d. Les trois quarts des patients analysés avaient un âge compris entre 24 et 49 ans
e. Le diagnostic de référence n'a pas pu être établi chez cinq patients, car l'appendicectomie n'a pas été réalisée (visualisation d'un appendice non inflammatoire lors de la laparoscopie)

**13** **À propos de la figure 4B, et des valeurs de performances diagnostiques de l'imagerie par résonance magnétique (MRI) pour les pathologies urgentes, il est correct de dire que**

a. Cette figure représente une courbe ROC
b. Elle permet de déterminer le seuil qui discrimine le mieux les patients avec pathologie urgente de ceux sans pathologie urgente
c. Les auteurs auraient pu présenter une aire sous la courbe, afin de comparer les performances diagnostiques des radiologues experts à celles des radiologues non experts
d. Plus le centre des croix est proche du coin supérieur gauche du graphique, meilleures sont les performances diagnostiques
e. La largeur de l'intervalle de confiance de la sensibilité, moins importante que celle de l'intervalle de confiance de la spécificité, peut s'expliquer par la valeur plus proche de 1 pour la sensibilité

**14** **Le(s) point(s) faible(s) de cette étude est(sont)**

a. L'absence de randomisation (lecture par experts contre lecture par non experts)
b. L'absence de tirage au sort des patients considérés pour l'analyse au sein des patients inclus dans l'étude
c. La nature composite du gold standard (reposant sur différents éléments cliniques et paracliniques)
d. La faible précision de la sensibilité de l'IRM lorsqu'elle est interprétée par les radiologues experts pour certains diagnostics, du au faible effectif de patients présentant effectivement ces diagnostics
e. La faible disponibilité actuelle de l'imagerie par résonance magnétique (MRI)

**15** **Le(s) point(s) fort(s) de cette étude est(sont)**

a. Son caractère multicentrique
b. Son caractère prospectif
c. Le tirage au sort des établissements participants au sein des établissements éligibles
d. L'interprétation de l'imagerie par résonance magnétique (MRI) en aveugle des résultats du scanner
e. La supériorité de la sensibilité et de la spécificité du scanner (CT) pour le diagnostic d'appendicite aiguë, par rapport à celles de l'imagerie par résonance magnétique (MRI) lorsqu'elle est interprétée par des radiologues non-experts



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 1

**Réponses exactes : B – E**

*Commentaires* Les arguments avancés par les auteurs pour justifier la réalisation de cette étude (*cf.* introduction) concernent les avantages de l'IRM par rapport au scanner : absence de rayonnement ionisant (B vraie), absence de néphropathie induite par les produits de contraste iodés (A fausse).

Les auteurs rapportent les résultats d'un certain nombre d'études ayant évalué l'intérêt diagnostique de l'IRM dans le cadre de l'appendicite (C fausse) ; la limite soulevée étant l'interprétation des IRM par des experts, et l'existence d'une seule étude ayant évalué l'accord inter-juge. Les auteurs soulignent également que peu de choses sont connues quant à l'exactitude diagnostique et la reproductibilité de l'IRM lorsque l'interprétation est réalisée par des non-experts (E vraie).

Les populations spécifiques des femmes enceintes et des enfants sont décrites, non pas par rapport à l'incidence de l'appendicite (D fausse), mais par rapport à l'intérêt de l'IRM comparativement au scanner dans ces sous-groupes.

## Question 2

**Réponses exactes : B – C**

*Commentaires* L'objectif principal de cette étude est présenté à la fin de l'introduction. Il comprend :

- la comparaison de la sensibilité et de la spécificité associées à l'interprétation de l'IRM par des radiologues experts *versus* des radiologues non experts pour le diagnostic d'appendicite, mais aussi pour d'autres pathologies urgentes (B vraie) ;
- l'évaluation de l'accord inter-juges entre interprétation de l'IRM par des radiologues experts et de radiologues non experts pour le diagnostic d'appendicite, mais aussi pour d'autres pathologies urgentes (C vraie).

Il n'est pas question dans cette étude de comparer les performances diagnostiques de l'IRM à celles du scanner, qui n'est pas le gold standard ici, ni un examen à l'étude (ADE fausses).

## Question 3

**Réponse exacte : D**

*Commentaires* La description de la procédure d'interprétation des IRM est décrite au début du paragraphe « 2.2 MRI interpretation ».

Chaque IRM est évaluée dans le cadre de trois lectures réalisées de façon indépendante :

- deux fois par deux radiologues non experts différents (parmi huit radiologues non experts participants) ;
- une fois par deux radiologues experts, qui en réalisent l'interprétation en consensus.

## Question 4

**Réponses exactes : B – E**

*Commentaires* La mise en place d'une formation pour les radiologues non experts a permis d'uniformiser et de standardiser la façon d'interpréter les IRM dans le groupe des non experts. Cela permet donc de limiter le biais de mesure, c'est-à-dire une erreur systématique dans la définition ou la mesure du critère de jugement (B vraie) et de diminuer la variabilité inter-juge chez les non experts (E vraie). Cela n'a pas d'influence sur le biais de sélection (A fausse) (définition des patients inclus).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Le biais de confusion ne se discute que dans le contexte des études étiologiques ou analytiques, ce qui n'est pas le cas ici (C fausse) (étude diagnostique).

La variabilité inter-interindividuelle des patients est un facteur indépendant des juges... La variabilité inter-individuelle des patients pourrait être influencée par le choix plus ou moins large des critères d'inclusion et de non-inclusion dans l'étude (D fausse).

### Question 5

**Réponse exacte : B**

*Commentaires* Dans cette étude, certains patients ont bu bénéficier d'une radiographie de l'abdomen sans préparation, mais ce n'était en aucun cas un examen réalisés systématiquement (absence de recommandation pour cet examen...) (A fausse).

L'IRM, examen à l'étude, a bien été réalisée pour tous les patients inclus et analysés (sept patients parmi les 230 patients inclus n'en ont pas bénéficié pour cause de problème technique ou de claustrophobie, mais ils n'ont pas été analysés pour les résultats de l'étude) (B vraie).

Il n'y a pas de randomisation dans le cadre de cette étude (C fausse).

Le scanner a pu être réalisé, mais ce n'était pas un examen systématique (réalisé en fonction des résultats de l'échographie) (D fausse).

Même si le tableau clinique était typique d'une appendicite aiguë, tous les patients inclus et analysés ont tous bénéficié d'une IRM (examen à l'étude) (E fausse).

### Question 6

**Réponses exactes : A – C – D – E**

*Commentaires* Le gold standard fait appel aux résultats issus des éléments suivants (*cf.* paragraphe 2.3) :

- données cliniques ;
- données d'imagerie (sauf l'IRM) ;
- chirurgie ;
- données anatomopathologiques ;
- suivi à trois mois.

L'IRM, qui est l'examen à l'étude, n'est pas un élément participant à la définition du gold standard dans cette étude (B fausse).

### Question 7

**Réponses exactes : A – C**

*Commentaires* Le nombre de sujets nécessaire, dans le cas d'un seul pourcentage qui est le cas ici (sensibilité), fait appel :

- au risque alpha (A vraie) ;
- à la valeur attendue du pourcentage évalué (C vraie) ;
- à la largeur de l'intervalle de confiance souhaitée.

Le risque bêta n'intervient pas lorsqu'on s'intéresse au nombre de sujets nécessaire pour la précision d'une estimation (pourcentage, moyenne...) (B fausse). Il n'intervient que lorsqu'on veut comparer deux valeurs (deux valeurs observées ou une valeur observée à une valeur théorique).

La différence attendue entre deux groupes n'intervient pas lorsqu'on s'intéresse au nombre de sujets nécessaire pour la précision d'une estimation (pourcentage, moyenne...) (D fausse). Elle n'intervient que lorsqu'on s'intéresse au nombre de sujets nécessaire pour mettre en évidence cette différence.

Ici les auteurs précisent qu'il s'agit d'un intervalle de confiance à 95 % et non à 90 % (E fausse).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 8

**Réponses exactes : B – C – D**

*Commentaires* Si les patients qui se présentent aux urgences pour suspicion d'une appendicite aiguë ont des caractéristiques différentes selon le moment de l'admission (jour de la semaine *versus* week-end *versus* jour férié, ou horaire de journée *versus* horaire de nuit), alors l'inclusion des patients uniquement dans certaines plages horaires a pu être à l'origine d'un biais de sélection (B vraie), et donc limiter la validité interne (D vraie) (existence de biais) et la validité externe de l'étude (C vraie) (capacité à extrapoler les résultats à la population cible).

Le biais de confusion ne se discute que dans le contexte des études étiologiques ou analytiques, ce qui n'est pas le cas ici (étude diagnostique) (A fausse).

La significativité des résultats obtenus dépendra notamment de la taille de l'effet observé, ou de l'effectif des patients analysés, qui sont indépendants des horaires et jours d'inclusion possibles (nombre de sujets nécessaire calculé : 230, ce qui correspond au nombre de patients initialement inclus, 7 ont été écartés pour cause de problème technique ou de claustrophobie) (E fausse).

### Question 9

**Réponses exactes : A – B – C – E**

*Commentaires* La sensibilité de l'IRM lorsqu'elle est interprétée par les radiologues non experts dans le cadre du diagnostic d'appendicite aiguë est de 89 % (probabilité que le test soit positif si le diagnostic est présent) (A vraie).

La sensibilité de l'IRM lorsqu'elle est interprétée par les radiologues experts dans le cadre du diagnostic d'appendicite aiguë est de 97 %, donc 1-sensibilité = 3 % (probabilité que le test soit négatif si le diagnostic est présent) (B vraie).

Les performances diagnostiques (sensibilité et spécificité) de l'IRM pour le diagnostic d'appendicite aiguë sont statistiquement meilleures lorsqu'elle est interprétée par des radiologues experts par rapport à des radiologues non experts (*cf.* résumé, toutes les p-values sont inférieures à la valeur seuil de 0,05) (C vraie).

L'indicateur qui permet d'évaluer la probabilité définie comme la probabilité que le diagnostic d'appendicite aiguë soit exact si un radiologue expert considère qu'un patient présente ce diagnostic est la valeur prédictive positive, qui est de 94 % et non de 96 % (valeur prédictive négative) (D fausse).

La prévalence de l'appendicite aiguë dans cette étude est bien de 52 % (*cf.* fin du paragraphe « 3. Results » : « … of which 117 patients with acute appendicitis (52 %…) ») (E vraie).

### Question 10

**Réponses exactes : A – C – D**

*Commentaires* Les performances diagnostiques (sensibilité et spécificité) de l'IRM pour le diagnostic d'une pathologie urgente sont statistiquement meilleures lorsqu'elle est interprétée par des radiologues experts par rapport à des radiologues non experts (*cf.* résumé, toutes les p-values sont inférieures à la valeur seuil de 0,05) (A vraie).

L'accord entre radiologues experts et radiologues non experts dans le cadre du diagnostic d'une pathologie urgente est évalué par le coefficient kappa, estimé à 0,63, ce qui qualifie l'accord de bon et non d'excellent (B fausse).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Puisque la sensibilité de l'IRM lorsqu'elle est interprétée par les radiologues experts pour le diagnostic de diverticulite aiguë vaut 1, c'est donc qu'il n'y a aucun faux négatif dans ce contexte (C vraie).

Puisque la sensibilité de l'IRM lorsqu'elle est interprétée par les radiologues experts pour le diagnostic d'occlusion intestinale vaut 1, alors le rapport de vraisemblance négatif (RVN = (1-Se)/Sp) dans ce contexte aurait bien été de 0, ce qui correspond en effet à un apport diagnostique très fort (grande probabilité de rejeter le diagnostic si le test est négatif).

L'intervalle de confiance à 95 % de la sensibilité pour le diagnostic des troubles des voies urinaires est asymétrique ; Se = 1,00 [0,65 – 1,00]. Il est normal que cet intervalle de confiance soit asymétrique, car la valeur maximale de la sensibilité est bornée à 1,00 (E fausse).

## Question 11

**Réponses exactes : A – B**

*Commentaires* Dans le cadre de cette étude, et plus spécifiquement dans le contexte de l'évaluation des performances diagnostiques de l'IRM lorsqu'elle est interprétée par des radiologues non-experts pour le diagnostic d'appendicite aiguë, le rapport de vraisemblance aurait pu être estimé (RVP = Se/(1-Sp)) (A vraie). L'indice de Youden aurait également pu être estimé (Se + Sp – 1) (B vraie).

Les autres éléments, aire sous la courbe ROC, graphique de Bland et Altman et courbe ROC, n'auraient pas pu être utilisés dans ce contexte spécifique. En effet, ils sont adaptés à l'évaluation des performances d'un test dont les résultats s'expriment sous forme de variable quantitative (dosage, taille tumorale, etc.) alors qu'ici le résultat de l'IRM s'exprime sous forme d'une variable qualitative (diagnostic présent/absent) (C, D, E fausses).

## Question 12

**Réponses exactes : B – C**

*Commentaires* Le nombre de patients initialement inclus pour cette étude est de 230, mais seuls 223 ont pu être analysés (sept n'ont pas bénéficié d'IRM pour cause de problème technique ou de claustrophobie) (A fausse).

Il y a bien trois patients avec un diagnostic d'occlusion intestinale (diagnostic d'après le gold standard, *cf.* table 2) (B vraie).

Il y a bien 330 patients non inclus car s'étant présentés hors des horaires définis pour l'étude (*cf.* paragraphe « 3. Results ») (C vraie).

L'intervalle interquartile de l'âge des patients analysés est de 24-49, ce qui veut donc dire que 25 % des patients sont âgés de moins de 24 ans, et que 25 % des patients sont âgés de plus de 49 ans, et que donc au total 50 % des patients sont âgés de 24 à 49 ans (D fausse).

Le diagnostic de référence a pu être établi chez l'ensemble des 223 patients analysés. Parmi les patients opérés, cinq d'entre eux n'ont pas subi d'appendicectomie car leur appendice était non inflammatoire. Pour autant, le diagnostic a tout de même pu être établi, étant donné qu'il ne repose pas uniquement sur les données anatomopathologiques (clinique, imagerie, suivi…) (E fausse).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 13

**Réponses exactes : D – E**

*Commentaires* Cette figure représente la sensibilité et la spécificité de l'IRM dans le cadre d'une pathologie urgente, lorsqu'elle est interprétée soit par des radiologues experts soit par des radiologues non experts, avec leur intervalle de confiance à 95 %, dans un espace ROC, mais il ne s'agit pas de courbe ROC (A fausse), adaptée à l'évaluation des performances d'un test dont les résultats s'expriment sous forme de variable quantitative (dosage, taille tumorale, etc.) alors qu'ici le résultat de l'IRM s'exprime sous forme d'une variable qualitative (diagnostic présent/absent). Cette figure ne permet donc pas de déterminer un meilleur seuil (un seuil n'a d'intérêt que lorsque le test s'exprime sous forme d'une variable quantitative) (B fausse). On ne peut donc pas non plus calculer d'aire sous la courbe (C fausse).

Plus le centre de la croix est proche du coin supérieur gauche du graphique, plus la sensibilité et la spécificité sont proches de 100 %, donc meilleures sont les performances diagnostiques (D vraie).

La largeur de l'intervalle de confiance d'un pourcentage de plusieurs paramètres, et notamment de son écart-type. Plus l'écart-type est grand, plus la largeur de l'intervalle de confiance est grande, tous les autres paramètres étant constants pas ailleurs. Cet écart-type est maximal pour un pourcentage de 50 %, et minimal pour un pourcentage de 0 % ou de 100 % (E vraie).

## Question 14

**Réponses exactes : D – E**

*Commentaires* La randomisation « lecture par les experts » contre « lecture par les non experts » n'auraient pas permis de répondre à une partie de l'objectif principal de cette étude (évaluation de l'accord entre non experts et experts, qui nécessite que toutes les IRM soient interprétées par les non experts ET les experts) (A fausse).

Parmi les patients inclus, les auteurs n'ont pas réalisé de tirage au sort pour déterminer un sous-groupe de patients à analyser. Cependant, ce n'est pas un point faible de l'étude. En effet, cela aurait amené à considérer un effectif plus faible de patients pour l'analyse, ce qui aurait été préjudiciable pour la précision des estimations. Il était donc approprié de considérer l'ensemble des patients inclus pour l'analyse (B fausse).

La nature composite du gold standard n'est pas une limite, car cela correspond à la réalité (C fausse). Tous les patients suspects d'appendicite aiguë ne sont pas opérés, il n'aurait donc pas été éthique de les opérer uniquement dans le cadre de cette étude diagnostique. L'important est que la définition du gold standard soit la plus claire possible, afin de limiter le biais de mesure.

La faible précision de la sensibilité de l'IRM pour certains diagnostics est en effet une limite à l'étude, qui est discutée par les auteurs (D vraie).

La faible disponibilité de l'IRM dans le contexte du diagnostic de pathologie urgente ou d'appendicite aiguë est une limite à l'étude (même si les performances diagnostiques avaient été parfaites, si l'IRM n'est pas un examen accessible, l'impact de l'étude s'en trouverait limité). Cette faible disponibilité limite l'utilité des résultats obtenus (E vraie).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 15**

**Réponses exactes : A – B – D**

*Commentaires* Le caractère multicentrique d'une étude est un point fort (améliore la capacité d'extrapolation des résultats car plus grande variabilité inter-sujet, recrutement plus rapide...) (A vraie). Son caractère prospectif est également un point fort (B vraie).

Le mode de sélection des établissements participants à cette étude n'est pas décrit dans l'article. Un tirage au sort aurait été un mode de sélection permettant de limiter le biais de sélection (ce qui aurait été positif), de même qu'une étude basée sur l'ensemble des établissements éligibles (étude exhaustive). Mais rien n'est précisé à ce sujet par les auteurs (C fausse).

L'évaluation des IRM se fait en aveugle des résultats du scanner, ce qui est un point fort de l'étude. En effet, l'IRM vise à remplacer le scanner, l'interprétation de l'IRM se fera donc à terme à l'aveugle des résultats du scanner (puisqu'il ne sera plus réalisé) (D vraie).

La supériorité des performances diagnostiques du scanner pour le diagnostic d'appendicite aiguë sur celle de l'IRM lorsqu'elle est interprétée par des radiologues non experts n'est pas un point fort de l'étude. En effet, cela limite l'impact de cette étude sur le changement des pratiques qui aurait pu en être attendu (E fausse).



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 13 Étude de pratique

Marie Herr

| | |
|---|---|
| **Titre de l'article** | **Client Factors Affect Provider Adherence to Clinical Guidelines during First Antenatal Care** |
| **Auteurs** | Amoakoh-Coleman M, Akua Agyepong I, Zuithoff NPA, Kayode GA, Grobbee DE, Klipstein-Grobusch K, Ansah EK |
| **Référence** | Plos One. 2016. 11(6) :e0157742. |
| **Lien vers l'article** | http://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0157542&type=printable |

**1** **Concernant le contexte de cette étude, les enjeux autour du premier examen prénatal**

a. Sont mineurs puisque les complications liées à la grossesse ont principalement lieu en fin de grossesse ou dans les 24 heures suivant l'accouchement
b. Sont la précocité de sa réalisation (au cours du premier trimestre de grossesse) afin d'identifier les situations à risque ou nécessitant un accompagnement particulier
c. Sont la réduction de la mortalité maternelle et infantile
d. Sont clairement présentés dans l'introduction et amènent à s'interroger sur le suivi des recommandations de pratiques concernant la réalisation du premier examen prénatal au Ghana
e. Portent notamment sur la prophylaxie contre le paludisme au Ghana

**2** **L'objectif de cette étude est**

a. De valider les recommandations de pratiques concernant le premier examen prénatal au Ghana (Safe Motherhood Protocol-SMP)
b. D'étudier le respect des recommandations SMP concernant le premier examen prénatal au Ghana
c. De comparer la réalité des pratiques aux recommandations SMP concernant le premier examen prénatal au Ghana
d. De décrire les facteurs influençant l'adhérence aux recommandations SMP concernant le premier examen prénatal parmi les caractéristiques des patientes et le type d'établissements de soins au Ghana
e. De réaliser une étude médico-économique du coût de la non-adhérence aux recommandations SMP au Ghana

**3** **Il s'agit d'une étude**

a. Randomisée
b. Transversale
c. De type enquête de pratiques
d. Observationnelle
e. Interventionnelle



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **Cette étude utilise les données collectées lors de la constitution d'une cohorte (données de « baseline »)**

a. Visant à évaluer l'impact de l'adhérence aux recommandations sur les complications de la grossesse
b. Dans laquelle on s'attend à observer deux fois moins de complications en cas d'adhérence incomplète aux recommandations de pratiques
c. Comprenant des sujets exposés (non-adhérence aux recommandations de pratiques) et des sujets non-exposés (adhérence)
d. Constituée d'un groupe de cas (non-adhérence aux recommandations) et d'un groupe de témoins (adhérence) appariés sur l'âge
e. Prévoyant un suivi des femmes et de leur nouveau-né jusqu'à 12 semaines après l'accouchement

**5** **La procédure d'échantillonnage**

a. Vise à inclure des femmes à partir de différents types d'établissements de soins (hôpitaux, polycliniques, centres de santé)
b. Inclut un tirage au sort des femmes à partir du registre de la population
c. Est dite en « grappes » (clusters), chaque établissement de soins correspondant à une grappe dans laquelle sont recrutées les femmes
d. Vise à inclure des femmes accouchant uniquement à l'hôpital
e. Inclut tous les établissements de soins de « the Greater Accra Region of Ghana »

**6** **Le recrutement des femmes**

a. A lieu lors de leur premier examen prénatal
b. A lieu en suites de couche
c. Concerne toutes les femmes satisfaisant aux critères de sélection entre décembre 2013 et avril 2014
d. N'est possible qu'après recueil de leur consentement
e. Concerne également les femmes n'ayant pas bénéficié d'un suivi médical en cours de grossesse

**7** **Le référentiel**

a. Se présente comme une grille de recueil à 13 items
b. S'appuie sur des protocoles internes à chaque établissement enquêté
c. S'appuie sur des recommandations faisant l'objet d'un fort consensus professionnel
d. Comporte des éléments concrets et observables permettant de porter un jugement valide sur la conformité de la pratique étudiée
e. Laisse une place au jugement de l'enquêteur sur la conformité des pratiques

**8** **La mesure de l'adhérence aux recommandations**

a. Permet de calculer un score d'adhérence sur 11
b. Repose sur l'évaluation de 13 critères à partir du dossier médical de la patiente
c. Est catégorisée *a posteriori* en 3 classes : adhérence complète/adhérence modérée/adhérence faible
d. Est rétrospective
e. Est prospective



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**9** **Concernant le tableau 1 :**

a. Il décrit les caractéristiques de la population d'étude
b. Il décrit la fréquence des 3 niveaux d'adhérence aux recommandations dans la population d'étude
c. Il décrit les facteurs associés aux 3 niveaux d'adhérence aux recommandations dans le cadre d'une analyse bivariée
d. Il décrit les facteurs associés aux 3 niveaux d'adhérence aux recommandations dans le cadre d'une analyse multivariée
e. Il montre que les 3 niveaux d'adhérence aux recommandations ont une prévalence du même ordre

**10** **Le score total d'adhérence tel qu'il est présenté dans le tableau 2**

a. Est exprimé sous forme de score moyen
b. Est maximum pour 48,1 % des femmes
c. A une valeur maximum de 11
d. A une valeur maximum de 13
e. A une bonne précision

**11** **Le modèle multivarié (tableau 3)**

a. Est présenté à l'aide de risques relatifs bruts et ajustés
b. Présente les facteurs associés à une adhérence totale aux recommandations au regard des critères du référentiel
c. Tient compte de la corrélation des mesures provenant d'un même établissement (cluster)
d. Montre que les femmes qui ont déjà eu un enfant ont plus de chances d'avoir un premier examen prénatal conforme aux recommandations par rapport aux femmes dont c'est le premier enfant
e. Montre que les complications au cours d'une précédente grossesse favorisent l'adhérence aux recommandations

**12** **Le caractère rétrospectif du recueil de données dans cette enquête de pratique**

a. Augmente le risque de données manquantes
b. Peut conduire à une sous-estimation de la bonne adhérence aux recommandations
c. Permet d'éviter un biais fréquent dans les enquêtes de pratiques qui est la modification des pratiques lorsque le soignant se sait observé
d. Augmente le risque de modification des pratiques du fait de la conscience de l'observation
e. Devrait conduire à parler d'une étude rétrospective plutôt que transversale



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**13** **Les forces de l'étude sont**

a. Le recrutement de toutes les femmes satisfaisant aux critères d'inclusion dans les centres participants
b. Son caractère multicentrique
c. L'inclusion d'établissements de soins publics et privés
d. La rigueur de l'analyse avec explication des données manquantes
e. L'analyse de l'influence du type de professionnel réalisant le premier entretien prénatal sur le suivi des recommandations

**14** **Les limites de l'étude sont**

a. Le manque de recommandations à l'issue de l'analyse des pratiques
b. L'absence d'hypothèses pour expliquer les écarts aux recommandations
c. L'absence de détails sur les types d'écarts aux recommandations observés dans les groupes à risque selon l'analyse multivariée (femmes ayant déjà eu des complications au cours d'une grossesse précédente ou vues dans des centres de santé généralistes)
d. Un poids trop important donné aux caractéristiques des patientes au regard des facteurs organisationnels dans le contexte d'une enquête de pratiques
e. De ne pas avoir considérer les facteurs associés à la réalisation ou à la non-réalisation du premier examen prénatal avant 5 mois de grossesse

**15** **Les auteurs ouvrent sur des perspectives d'études qualitatives pour mieux comprendre les facteurs expliquant la non-adhérence aux pratiques du point de vue des professionnels de santé. Une étude qualitative est**

a. Une étude de très bonne qualité
b. Une étude reposant sur des entretiens individuels ou collectifs
c. Une étude utilisant des méthodes relevant de la sociologie
d. Une étude ayant une bonne validité interne
e. Une étude reposant sur la passation de questionnaires standardisés à des professionnels de santé



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 1

**Réponses exactes : B – C – D – E**

*Commentaires* Même si les complications liées à la grossesse surviennent principalement en fin de grossesse ou dans les 24 heures suivant l'accouchement, les auteurs insistent sur l'importance de suivre au plus tôt la grossesse et d'identifier les situations à risque dans le cadre du premier examen prénatal dans un objectif de prévention de ces complications en fin de grossesse.

### Question 2

**Réponses exactes : B – C – D**

*Commentaires* Cette étude vise à étudier les pratiques professionnelles dans le cadre du premier examen prénatal. Il s'agit d'une enquête de pratiques de type audit dans laquelle on compare la réalité des pratiques aux recommandations en vigueur au moment de l'enquête afin de mettre en évidence ce qui pourrait être amélioré.

### Question 3

**Réponses exactes : B – C – D**

**Indispensable : C**

*Commentaires* Cette étude est observationnelle : il s'agit d'observer l'adhérence aux recommandations, qui n'est pas imposée mais observée par les investigateurs. L'emploi du terme « arms » dans le calcul du nombre de sujets nécessaires peut induire en erreur ; ces bras correspondent aux groupes exposés et non exposés (exposition = adhérence aux recommandations) pour l'analyse de l'influence de l'adhérence aux recommandations de pratiques sur le risque de complications. L'étude est transversale car elle évalue l'état des pratiques à un moment donné.

### Question 4

**Réponses exactes : A – C**

*Commentaires* L'hypothèse utilisée pour ce calcul du nombre de sujets nécessaires est celle d'un taux de complications deux fois supérieur en cas d'adhérence incomplète aux recommandations de pratiques. Il ne s'agit pas d'une étude cas-témoins mais d'une étude de cohorte dans laquelle on compare les évènements de santé d'un groupe exposé à un groupe non-exposé. Le suivi des femmes et de leur nouveau-né était limité à 6 semaines après l'accouchement.

### Question 5

**Réponses exactes : A – C**

*Commentaires* La procédure d'échantillonnage concerne les établissements de santé dans lesquelles les patientes ont été recrutées. En plus de l'hôpital régional, les auteurs ont sélectionné par tirage au sort des établissements de soins (hôpitaux, polycliniques, centres de santé) de différentes provinces dotées ou non d'un hôpital publique. Toutes les femmes satisfaisant aux critères de sélection et acceptant l'étude ont ensuite été incluses dans chacun des centres d'enquête, sans tirage au sort à ce niveau.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**Question 6**

**Réponses exactes : B – C – D**

**Inadmissible : E**

*Commentaires* Les femmes n'ayant pas eu leur premier examen prénatal avant 5 mois de grossesse et dans le centre où elles ont accouché ne rentrent pas dans le périmètre de l'étude puisqu'il s'agit de collecter des données sur le premier examen prénatal et de les mettre en lien avec les éventuelles complications ultérieures.

**Question 7**

**Réponses exactes : A – C – D**

*Commentaires* L'évaluation des pratiques professionnelles se fait le plus souvent par comparaison des pratiques observées à une liste de critères de « bonne pratique » faisant l'objet d'un consensus professionnel fort (conférence de consensus, recommandations nationales ou internationales...). Ici, le référentiel utilise les recommandations nationales de prise en charge, ce qui semble pertinent dans ce contexte. Les critères d'évaluation sont soumis aux mêmes exigences que les critères de jugement des essais cliniques : ils doivent être simples, valides et observables en pratique. Leur évaluation ne doit pas laisser de place au jugement de l'enquêteur.

**Question 8**

**Réponses exactes : A – B – C – D**

*Commentaires* La mesure de l'adhérence repose sur la présence ou l'absence dans le dossier médical de 13 éléments indispensables à la bonne réalisation du premier examen prénatal, dont 11 éléments obligatoires et 2 optionnels dépendant de l'âge gestationnel. Chaque élément vaut 1 point dans le calcul du score. Le score d'adhérence est une variable quantitative, catégorisée au moment de l'analyse en variable semi-quantitative avec 3 niveaux d'adhérence aux recommandations. La mesure est rétrospective car elle implique d'aller rechercher l'information dans le dossier médical des femmes incluses en suite de couche.

**Question 9**

**Réponses exactes : A – B – C**

*Commentaires* Ce premier tableau fait office de tableau descriptif de la population et d'analyse bivariée des caractéristiques associées aux 3 niveaux d'adhérence aux recommandations.

**Question 10**

**Réponses exactes : A – B – D – E**

*Commentaires* La proposition C est fausse car il est impossible d'avoir une valeur maximum de 11 et une moyenne de 11,73. Bien que les auteurs annoncent dans le paragraphe statistique de la méthode un score sur 11, le tableau 2 présente la moyenne du score sur 13. L'intervalle de confiance de la mesure est étroit donc la mesure est précise.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 11

**Réponses exactes : B – C – D**

**Inadmissible : E**

*Commentaires* Le modèle multivarié est un modèle de régression logistique, ses résultats sont donc exprimés sous forme d'Odds Ratio (OR) et non de risque relatif. Le risque relatif est un rapport d'incidence qui ne peut être calculé que lorsque les données sont longitudinales, ce qui n'est pas le cas dans cette étude transversale. La modélisation tient compte de la dépendance ou « ressemblance » entre mesures provenant d'un même établissement (potential cluster effect). Les OR significativement inférieurs à 1 indiquent les facteurs allant à l'encontre de l'adhérence aux recommandations, soit les facteurs de risque de non-adhérence.

### Question 12

**Réponses exactes : A – B – C**

*Commentaires* Le recueil des données concernant les éléments du premier examen prénatal est rétrospectif à partir du dossier médical des patientes. En utilisant ce mode de recueil des données, on ne risque pas de modifier les pratiques par rapport à une étude où le recueil des données est concomitant à la réalisation des soins et où le professionnel peut modifier ses habitudes car il a conscience d'être observé (effet Hawthorne). Cependant, ce type de recueil de données est plus à risque de données manquantes. Le poids, par exemple, peut avoir été mesuré mais pas reporté dans le dossier de la patiente au moment de l'examen. L'item mesure du poids est alors coté 0. In fine, cela peut conduire à sous-estimer l'adhérence aux recommandations. Le caractère transversal de l'étude n'est pas remis en cause par le recueil de données rétrospectif. Même s'il existe un décalage de quelques mois entre le premier examen prénatal et le recrutement, l'étude est dite transversale car elle décrit la prise en charge d'une population de femmes au moment de la réalisation de l'étude. Une étude rétrospective consisterait à étudier des dossiers anciens, de patientes ayant accouché au cours des années précédentes.

### Question 13

**Réponses exactes : A – B**

*Commentaires* Le recrutement de toutes les femmes éligibles dans chaque établissement est une force de cette étude car cela limite les erreurs d'échantillonnage et les biais de sélection. Même si différents types d'établissements de soins sont inclus dans cette étude (hôpitaux, polycliniques, centres de santé), l'absence d'établissements de soins privés constitue une limite de ce travail (résultats moins généralisables). La gestion des données manquantes, dont on s'attend à ce qu'elles ne soient pas négligeables compte tenu des modalités du recueil des données (rétrospectif à partir de dossiers médicaux) n'est pas développée dans l'article. Une analyse rigoureuse présenterait la fréquence des données manquantes et leurs facteurs associés pour que le lecteur puisse juger de leur possible influence sur les résultats de l'étude.

### Question 14

**Réponses exactes : A – C – D – E**

**Indispensable : D**

*Commentaires* Si les auteurs formulent des hypothèses sur les raisons qui peuvent expliquer les écarts aux recommandations, notamment pour expliquer pourquoi les femmes ayant déjà eu des complications en cours de grossesse sont plus à risque d'avoir un examen incomplet, ils ne reviennent pas sur les 13 items de l'évaluation dans les groupes identifiés comme à risque par l'analyse multivariée. Cela pourrait pourtant apporter des éléments explicatifs sur des



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

écarts inattendus *a priori* et déboucher sur des pistes d'amélioration concrètes. En termes de prévention, les femmes ne bénéficiant pas d'un examen prénatal avant le 5e mois de grossesse sont une cible prioritaire. Celles-ci sont occultées dans cette enquête. Même s'il s'agit d'une autre problématique qui est l'accès à l'examen prénatal, il aurait été intéressant d'avoir des informations sur la proportion de femmes n'ayant pas eu d'examen prénatal avant 5 mois de grossesse dans cette étude afin d'avoir une vision plus globale de la situation. Enfin et surtout, la place donnée aux caractéristiques des patientes dans la recherche des déterminants de l'adhérence aux recommandations semble démesurée au regard de l'attention portée aux facteurs organisationnels. Cette limite est reconnue par les auteurs mais peu justifiée. On peut pourtant supposer que les données sur la qualification de la personne ayant réalisé l'examen prénatal et sur les caractéristiques des établissements enquêtés (activité, ressources humaines, organisation du travail...) sont relativement accessibles.

■ **Question 15**

**Réponses exactes : B – C**

*Commentaires* Les études qualitatives peuvent aider à approfondir et à donner du sens aux résultats d'études quantitatives ou contribuer à les préparer. Dans ce cadre, la réalisation d'entretiens individuels ou collectifs (focus group) avec les personnes intervenant dans la réalisation des examens prénatals pourrait faire émerger des habitudes, des logiques de soins ou des contraintes difficilement évaluables par des questionnaires standardisés. Les entretiens sont littéralement retranscrits et analysés selon des méthodes empruntées à la sociologie et à la psychologie.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Annales /ECN blanc

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 1 Épreuve de LCA 2016

Hervé Devilliers, Vanessa Cottet, Julie Berbis, Moustapha Dramé, Christine Binquet

| | |
|---|---|
| **Titre de l'article** | **Maternal Consumption of Coffee and Caffeine-containing beverages and oral clefts : a population based case-control study in Norway** |
| **Auteurs** | Johansen AM, Wilcox AJ, Lie RT, Andersen LF, and Drevon C |
| **Référence** | Am J Epidemiol. 2009 ; 169(10):1216-22 |
| **Lien vers l'article** | https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2727209/ |

**1** **Il s'agit d'une étude**

a. Randomisée
b. Rétrospective
c. Observationnelle
d. Étiologique
e. Prospective

**2** **Dans cette étude, les cas sont les nouveau-nés**

a. Présentant une fente labio-palatine à la naissance, opérés dans les centres hospitaliers universitaires d'Oslo et de Bergen
b. De mère qui consommaient plus de 3 tasses de café par jour au premier trimestre de grossesse
c. Dont la mère a accouché dans les maternités des centres hospitaliers universitaires d'Oslo ou de Bergen
d. Indemnes de fente labio-palatine à l'issue de la grossesse
e. Décédés après la naissance

**3** **Dans cet article, les facteurs d'exposition d'intérêt incluent**

a. La consommation de thé au cours du troisième trimestre de grossesse
b. La supplémentation vitaminique, au cours du premier trimestre de la grossesse
c. La consommation de boissons contenant de la caféine au cours du premier trimestre de grossesse
d. Le tabagisme au cours du premier trimestre de grossesse
e. La consommation de café au cours du premier trimestre de grossesse



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **L'information sur la consommation quotidienne de café au cours du premier trimestre de grossesse**

a. A été de nature déclarative
b. A été recueillie de manière prospective
c. A été extraite de la base de données du registre national des naissances
d. A été quantifiée en nombre de tasses
e. A été recueillie uniquement chez les mères de nouveau-nés avec une fente labio-palatine

**5** **Le choix d'une étude de type cas-témoin a été guidé par le fait ou les faits suivants**

a. La prévalence de la fente labio-palatine à la naissance est faible
b. Ce type d'étude permet de montrer une relation causale
c. Ce type d'étude ne présente généralement pas de biais
d. La consommation de café au cours du premier trimestre de grossesse est fréquente
e. Le niveau d'exposition aux facteurs de risque étudiés peut être déterminé précisément

**6** **Dans cette étude, approximativement deux témoins ont été recrutés pour un cas de fente labiale (avec ou sans fente palatine) dans le but**

a. De limiter un potentiel biais de sélection lié aux sujets perdus de vue en cours de suivi
b. De calculer des risques relatifs
c. De réduire la largeur de l'intervalle de confiance des odds ratios estimés
d. D'augmenter les chances de mettre en évidence une différence entre les cas et les témoins
e. De procéder à un appariement sur les facteurs de confusion potentiels

**7** **D'après le tableau 1**

a. Avant la grossesse, 75 % des femmes consommaient au moins une tasse de café par jour
b. La fréquence du tabagisme maternel actif au cours du premier trimestre de grossesse était de 31,9 % dans le groupe témoin
c. 79,5 % des mères travaillant durant le premier trimestre de grossesse ont donné naissance à un nouveau-né avec une fente palatine seule (FPS)
d. 50 % des mères du groupe témoin consommaient au moins deux tasses de café quotidiennement pendant le premier trimestre de grossesse
e. 573 cas de fente labio-palatine (FLP) ou fente palatine seule (FPS) ont été inclus au total dans cette étude



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**8** **Pourquoi le tabagisme actif est-il un facteur de confusion potentiel de l'association observée entre la consommation de café au cours du premier trimestre de grossesse et le diagnostic de fente labio-palatine à la naissance dans cette étude ?**

a. Car la prévalence du tabagisme actif au premier trimestre de grossesse diffère entre les cas et les témoins dans cette étude
b. Car le tabagisme est un facteur de risque des fentes labio-palatines identifié dans la littérature
c. Car il est possible que les femmes fumeuses consomment plus de café que les non fumeuses
d. Car l'information sur le tabagisme a été recueillie rétrospectivement par questionnaire dans cette étude
e. Car le tabagisme passif et l'absence de tabagisme ont été regroupés au sein d'une seule et même catégorie dans cette étude

**9** **Quelle(s) stratégie(s) ont utilisée(s) les auteurs pour s'affranchir de l'effet du facteur de confusion « tabagisme actif au cours du premier trimestre de grossesse » dans cette étude ?**

a. L'estimation d'odds ratios bruts
b. Un test de tendance sur la consommation de café
c. L'appariement de témoins à chaque cas
d. La restriction aux mères non-fumeuses
e. L'ajustement à l'aide d'un modèle de régression logistique multivariée non-conditionnelle

**10** **En analyse multivariée (Tableau 2), les auteurs indiquent qu'il existe une relation croissante entre l'odds ratio de fente labiale (avec ou sans fente palatine, FPL) et la consommation de café codée en tasses par jour. Sur quelle(s) preuve(s) repose cette affirmation ?**

a. Le recouvrement partiel des intervalles de confiance à 95 % des odds ratios ajustés pour les catégories > 0- < 3 tasses et 3 tasses et plus, respectivement
b. L'exclusion de l'échantillon d'analyse des mères ne consommant pas de café pendant le premier trimestre de grossesse
c. L'odds ratio de fente labiale chez les femmes consommant 3 tasses par jour ou plus de café par rapport aux femmes consommant moins de 3 tasses par jour est de 1,59 ($IC_{95\ \%}$ 1,052,39)
d. L'estimation ponctuelle de l'odds ratio ajusté égale, respectivement, à 1,00 pour 0 tasse, 1,39 pour > 0- < 3 tasses, et 1,59 pour 3 tasses et plus
e. Le degré de signification (valeur de p) du test de tendance (soit 0,013)



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**11** **Le risque de conclure à tort à la responsabilité de la consommation de café au cours du premier trimestre de la grossesse dans cette étude peut s'expliquer par**

a. Le risque d'erreur statistique de première espèce (alpha)
b. Le caractère rétrospectif du recueil des données
c. L'existence d'une relation entre la dose de caféine ingérée quotidiennement et l'odds ratio de fente labiale
d. L'existence de facteurs de confusion résiduels méconnus
e. Un défaut de puissance statistique

**12** **Sur la base des informations données dans cet article, quel(s) facteur(s) de confusion potentiel(s) supplémentaire(s) les auteurs auraient-ils pu prendre en compte ?**

a. Les antécédents familiaux
b. La consommation de café au deuxième trimestre
c. L'homocystéinémie chez la mère
d. L'âge gestationnel
e. L'activité physique

**13** **Quels sont les biais potentiels de cette étude ?**

a. Biais de sélection des populations étudiées
b. Biais lié à la présence de facteurs de confusion
c. Biais de classement sur les critères définissant les cas
d. Biais de classement sur l'exposition, lié à l'absence de souvenir
e. Biais d'attrition des populations étudiées

**14** **Citez le(s) argument(s) en faveur de la nature causale de l'association, retrouvée dans cette étude, entre la consommation de café pendant le premier trimestre de grossesse et le risque accru de fente labiale (avec ou sans fente palatine, FLP) à la naissance**

a. L'ajustement sur les facteurs de confusion connus
b. L'existence d'une relation dose-effet statistiquement significative
c. La séquence temporelle entre l'exposition et l'événement de santé étudié
d. Le risque de fente labio-palatine induit par l'exposition à une dose unique élevée de caféine chez l'animal
e. L'absence d'association similaire entre la consommation de café et la fente palatine seule

**15** **Parmi les implications potentielles de cette étude pour la prévention des fentes labio-palatines, il paraît pertinent**

a. De réaliser une étude randomisée portant sur la consommation de thé *versus* café chez les femmes enceintes
b. D'inciter les femmes enceintes à consommer plus de trois tasses de thé par jour pendant le premier trimestre de grossesse
c. De renforcer la surveillance échographique morphologique de l'extrémité céphalique chez les femmes enceintes consommant plus de trois tasses de café par jour
d. De conseiller d'éviter les boissons contenant de la caféine au cours du premier trimestre de grossesse
e. D'informer les femmes enceintes concernées sur l'absence de preuve de l'innocuité d'une consommation de café supérieure à trois tasses par jour



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

## Question 1

**Réponses exactes : B – C – D**

*Commentaires* Cet article repose sur une étude cas-témoin. Il s'agit donc d'une étude observationnelle (on ne change pas la prise en charge des sujets) et rétrospective (les fentes labiales et/ou palatines, qui correspondent aux pathologies étudiées, est présente avant le début de l'étude – les inclusions interviennent dans les premières semaines de vie). Il s'agit clairement ici d'une étude étiologique car les auteurs cherchent les expositions (notamment café, thé ou boissons caféinées) pouvant être en cause dans la survenue de fentes labiales (FL) et/ou palatines (FP). Ce type d'étude rentre dans la catégorie des études analytiques qui aurait pu être l'un des mots-clés proposés. Même si les témoins ont été tirés au sort, on ne peut pas parler d'étude randomisée. Ce terme n'est utilisé que pour l'allocation aléatoire des sujets inclus dans différents groupes d'intervention.

## Question 2

**Réponse exacte : A**

*Commentaires* La seule réponse exacte est l'item A. La définition des cas est abordée au début du paragraphe méthode. Il est précisé que les enfants avec une fente labiale et/ou palatine, qui constituent les cas, sont pris en charge uniquement dans 2 centres en Norvège. Ces deux centres participent tous les deux à l'étude permettant d'obtenir une population représentative des cas. Cela permet aux auteurs de s'affranchir du lieu d'accouchement des mères (item C faux). L'item B est bien sûr faux, car si l'inclusion était conditionnée par le fait d'avoir été exposé à une consommation de café, de thé ou de boisson caféinées, on ne pourrait plus comparer le pourcentage d'exposition chez les cas et les témoins. L'item D correspondrait à des témoins et non à des cas. Enfin l'item E est également faux puisque ce critère était un critère de non-inclusion, pour des questions éthiques.

## Question 3

**Réponses exactes : C – E**

*Commentaires* Les facteurs d'exposition d'intérêt sont détaillés dans les parties « data collection », « methods », mais aussi dans le tableau 1. La question vous incite à cibler non pas l'ensemble des facteurs d'exposition pris en compte dans l'étude et susceptible d'influencer la survenue de fentes labio-palatines, mais bien ceux qui font l'objet de la recherche soit spécifiquement la consommation de café, de thé et de boissons caféinées au cours du 1er trimestre de grossesse (item A faux). Les autres facteurs d'exposition déjà identifiés dans la littérature sont considérés comme des facteurs de confusion (c'est le cas du tabagisme, et de la supplémentation vitaminique – items B et D faux). Ces facteurs sont pris en compte dans les analyses multivariées en tant que facteurs d'ajustement. On retrouve une description précise des ajustements réalisés dans les notes de bas de tableau du tableau 2, sans que les OR correspondant soient présentés spécifiquement.

## Question 4

**Réponses exactes : A – D**

*Commentaires* Les consommations de café ont été recueillies après l'inclusion dans les premières semaines de vie des enfants que ce soit dans le groupe des cas ou des témoins (questionnaire remplis en moyenne 14 à 15 semaines après l'accouchement. Le recueil était donc bien après l'exposition recherchée (1er trimestre) et est de nature rétrospective. Les femmes devaient remplir un questionnaire de 32 pages détaillant leur consommation de boissons contenant du café.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

L'information était donc déclarative et non extraite d'une base de données. L'exposition était bien mesurée en nombre de tasse. Enfin l'item E est clairement faux, car sans recueil de l'exposition chez les cas et les témoins il ne serait pas possible de réaliser la comparaison souhaitée.

### Question 5

**Réponses exactes : A – D**

*Commentaires* Les études cas-témoins sont les études choix pour mettre en relation des expositions fréquentes et des maladies rares (item A et D vrais). La prévalence des fentes labio-palatines est un peu au-dessus du seuil définissant une maladie rare (2/1 000), mais permet de considérer que dans cette étude les OR sont de bonnes estimations des Risques Relatifs. En général on considère que c'est le cas dès que la prévalence dans la population (et non dans l'échantillon) est inférieure à 5 %. Ce type d'étude ne permet pas de montrer une relation causale, puisqu'elle est particulièrement sujette aux biais (item B et C faux). Par définition rétrospective, il implique souvent un biais de mémorisation source de biais de classement vis-à-vis de l'exposition (item E faux), qui peut être en plus différent chez les cas et les témoins entre les deux groupes, les mères des cas pouvant être amenées à minimiser ou maximiser leurs déclarations. Les études cas-témoins sont également sujettes à des biais de sélection des cas et des témoins ne serait-ce que par l'acceptation de répondre ou non aux questionnaires de l'étude. Par contre il peut participer au faisceau d'arguments qui vont permettre de conclure à une relation causale, notamment quand il met en évidence une relation « dose-effet ».

### Question 6

**Réponses exactes : C – D**

*Commentaires* Le choix d'inclure deux témoins par cas a pour objectif d'augmenter la puissance de l'analyse statistique à nombre de cas constants. En pratique cela permet d'augmenter les chances de conclure à une association entre exposition et maladie si cette association existe (item D vrai). Cela permet aussi de réduire l'étendue des intervalles de confiance et ainsi d'augmenter la précision des estimations. Il est possible d'inclure plus de 2 témoins par cas, mais au-delà de 4 témoins par cas, le gain de puissance devient limité. Par ailleurs, en cas d'étude appariées, plus le nombre de témoins par cas est élevé, plus il est difficile de procéder à un appariement c'est-à-dire à la sélection de témoins présentant des caractéristiques similaires à chacun des cas (item E faux). Un Risque Relatif (RR) correspond au ratio de la proportion de cas chez les exposés sur la proportion de cas chez les non exposés. Une étude cas-témoin ne permet donc jamais de calculer de RR, puisque la proportion cas et de témoins est imposée par le design de l'étude. L'item A est bien sûr faux puisque l'étude cas-témoin est réalisée en aval de l'exposition et de la maladie.

### Question 7

**Réponses exactes : B – E**

*Commentaires* La description des consommations de café, de thé, et d'alcool sont décrites dans le tableau 1 en termes de médiane (et d'intervalle interquartile). La médiane est la valeur séparant en deux parties égales la population : 50 % ont une consommation supérieure à cette valeur, 50 % ont une consommation inférieure. On parle aussi de 2e quartile. L'intervalle interquartile a pour borne intérieure la valeur seuil du 1er quartile – ou 25e percentile – et pour borne supérieure, la valeur seuil du 3e quartile – ou 75e percentile. En pratique, par exemple, les valeurs rapportées pour la consommation de café par jour dans le groupe témoin, indiquent que :

- 50 % des femmes consommaient moins de 3 tasses par jour avant la grossesse



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

- 25 % en consommaient moins d'une tasse par jour (et donc 75 % consommaient plus d'une tasse par jour)
- 25 % avaient une consommation supérieure à 4,3 tasses par jour (ce qui revient à dire que 75 % consommaient moins de 4,3 tasses par jour)
- Ceci indique que 50 % des femmes consommaient entre 1 tasse et 4,3 tasses par jour avant la grossesse.

Si on examine l'ensemble des données pour la période avant la grossesse, tous les intervalles interquartiles ont comme borne inférieur la valeur 1 indiquant que 75 % consommaient au moins une tasse de café par jour, mais les notes de bas de tableau précise que ces valeurs correspondent à la consommation des consommatrices de café et non de la totalité des femmes dans chaque groupe (item A faux). L'item D est du coup également faux, 50 % des femmes du groupe témoin, consommatrices de café consommaient au moins 2 tasses par jour. Pour les items B et C il faut bien repérer comment ont été calculé les pourcentages. Les pourcentages de chaque caractéristique sont calculés en prenant comme dénominateur l'effectif du groupe de femmes (fente labio-palatine, fente palatine seule, groupe témoin). L'item est B est donc vrai, mais l'item C est faux. Le tableau ne présente pas la fréquence des fentes palatines seules parmi les femmes travaillant au 1er trimestre, mais le pourcentage de femmes travaillant au 1er trimestre dans le groupe des femmes ayant donné naissance à un enfant avec une fente palatine seule. Enfin pour l'item E, il suffit d'additionner les effectifs des deux groupes de cas.

### Question 8

**Réponses exactes : B – C**

*Commentaires* Un facteur de confusion est un facteur lié à la fois à la survenue de la maladie et à l'exposition. Le fait que la question précise facteur de confusion « potentiel » indique que le raisonnement se positionne *a priori*, dans le cadre de la planification de l'étude. De ce fait, l'item A est faux car il s'appuie sur les résultats de l'étude et non sur une donnée préalable. Les items B et C concernent en revanche bien des données *a priori* et participent tous deux à la justification du rôle confondant du tabagisme : l'item B met en avant son lien avec la maladie, et l'item C avec l'exposition d'intérêt. Les items D et E sont sans rapport avec la question.

### Question 9

**Réponses exactes : D – E**

*Commentaires* De façon générale, il existe plusieurs manières de limiter l'impact des facteurs de confusion :

- dans les études expérimentales, on peut allouer de façon aléatoire l'exposition (principe de la randomisation). Cette procédure permet que les caractéristiques (incluant celles représentant des facteurs de confusion) des groupes exposés et non exposés soient comparables ;
- dans les études observationnelles telles que cette étude cas-témoin, on peut :
  - apparier les cas et les témoins sur le ou les facteurs de confusion. Cet appariement permet d'obtenir des groupes comparables pour ce ou ces ce facteurs, mais on ne peut plus étudier leur impact sur la maladie. Plus le nombre de facteurs sera important plus il sera difficile de trouver des témoins pour chacun des cas. Ici les auteurs n'ont pas appliqué cette procédure (item C faux) ;
  - ajuster les analyses permettant d'estimer l'association entre exposition et maladie sur les facteurs de confusion ; on parle d'analyse multivariée (ici régression logistique) par contraste avec les analyses bivariées (ou analyses brutes) incluant uniquement



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

exposition et maladie. Dans ce dernier cas on obtient un OR brut (« crude » en anglais) qui ne tient pas compte du facteur de confusion (item A faux). Ces analyses permettent d'estimer l'effet propre de l'exposition d'intérêt sur la survenue de la maladie étudiée. On obtient alors un OR ajusté (item E vrai) ;

- limiter l'analyse à un seul profil de sujets (par exemple analyse limitée aux non-fumeuses), ce qui a été fait ici ; cela peut être une analyse de sensibilité comme ici, ou conduire à ne mener la totalité de l'étude dans ce seul groupe (item D vrai) ;
- réaliser une analyse dans chaque sous-groupe défini par le facteur de confusion. On parle alors d'analyse « stratifiée ». Cette technique ne permet pas d'étudier l'impact du facteur de confusion sur la maladie.

L'item B est faux, car un test de tendance permet de tester si on a par exemple une augmentation de la probabilité d'observer la maladie dans des catégories d'exposition croissante mais ne limite en aucun cas l'effet des facteurs de confusion.

### Question 10

**Réponses exactes : D – E**

*Commentaires* Cette question cible les analyses ajustées réalisées pour les fentes labiales avec ou sans fente palatine. Trois catégories de consommation de café sont prises en compte : 0, 0-3, > 3 tasses. L'OR mentionné pour chaque catégorie correspond au risque de fente labio-palatine chez les enfants des femmes ayant ce niveau de consommation par rapport au risque chez les enfants des femmes ne consommation pas de café, qui constituent la catégorie de référence. Ces femmes ne sont bien sûr pas exclues, l'item B est faux.

Les OR sont supérieurs à 1 avec des intervalles de confiance ne comprenant pas la valeur à 1. Cela signifie que pour les deux catégories de consommation de café le risque de fentes est significativement plus important que pour les femmes n'ayant pas consommé de café. Par contre ils ne permettent pas de comparer directement le risque de ces deux catégories (item C faux). Lorsque la catégorie de référence est mentionnée dans un tableau de résultat, l'OR correspondant est par définition égal à 1 et cette valeur est mentionnée sans intervalle de confiance. Ici, la catégorie de référence est notée en bas de tableau mais elle est souvent implicite et il faut la déduire des autres éléments de l'article.

Ici, par comparaison aux femmes n'ayant pas consommé de café dont le risque est par définition égal à 1 puisqu'il s'agit de la catégorie de référence, l'OR passe de 1,39 pour les femmes ayant consommé moins de 3 tasses par jour, à 1,59 pour celles consommant plus de 3 tasses par jours. Ces estimations sont en faveur d'une relation croissante entre consommation de café et fente labio-palatine. Le test de tendance (p for trend = 0,013) confirme que cette relation croissante est significative. On parle alors de relations « dose-effet ».

### Question 11

**Réponses exactes : A – B – D**

*Commentaires* Le risque de conclure à tort, peut relever du hasard (item A vrai) : les tests peuvent être significatifs à tort du fait de fluctuations d'échantillonnage et on peut observer des associations qui en réalité n'existent pas ; la p-value permet de mesurer la probabilité d'observer une association aussi élevée dans l'échantillon sous l'hypothèse que l'OR est en fait égal à 1 et permet ainsi de disposer d'une estimation du risque de conclure à tort. Plus la valeur de p sera proche de 0, plus ce risque sera limité. Ici p = 0,013 et indique que la probabilité de conclure à tort est de l'ordre de 1 %. Au seuil de 0,05 on peut donc conclure à une association significative avec risque limité de conclure à tort ; il s'agit donc d'un argument allant plutôt dans le sens de



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

la réalité de l'association (item C faux). Une autre raison est la présence de biais de classement, surtout s'ils sont différentiels, ce qui est le risque dans une étude cas-témoin, par définition retrospective (item B vrai). En effet, il n'est pas impossible que les mères d'enfants nés avec une fente labio-palatine aient maximisé leur consommation de café par rapport aux femmes du groupe témoin, car elles se sont probablement interrogées sur les causes de ces fentes et surestimé leur consommation. Dans ce cas, cela peut conduire à une surestimation du différentiel d'exposition entre les deux groupes et ainsi contribué à surestimer l'association entre consommation de café et fente labio-palatine. La dernière raison pourrait être l'existence de facteurs de confusion non pris en compte dans l'étude. Les auteurs ont ajusté leurs analyses sur la majorité des facteurs envisageables, mais il n'est pas impossible que d'autres facteurs puissent jouer (item D). Seule une randomisation (bien sûr inenvisageable ici) pourrait contourner ce problème en équilibrant les groupes de façon globale.

### Question 12

**Réponse exacte : A**

*Commentaires* Les antécédents familiaux, signalés comme facteur de risque potentiel n'ont pas été pris en compte (item A vrai). La consommation de café au 2e trimestre intervient après la constitution des lésions (8-12 SA), elle n'est donc pas un facteur de confusion (idem pour l'âge gestationnel). La relation avec l'activité physique n'est pas évoquée. L'homocystéinémie était impossible de la prendre en compte dans le cadre de cette étude rétrospective. Cependant, ce facteur pourrait s'intégrer dans la chaîne causale entre consommation de café et fente labio-palatine. Il ne s'agit plus alors d'un facteur de confusion, mais plutôt d'un facteur intermédiaire. L'ajustement des analyses sur ce type de facteur ne permet pas de s'affranchir complètement de leur influence

### Question 13

**Réponses exactes : A – B – D**

*Commentaires* Par construction, une étude cas-témoin est susceptible aux 3 grands types de biais, quelle que soit la qualité de sa conception.

1. Biais de sélection : ce biais intervient lorsque les cas et/ou les témoins ne sont pas représentatifs de la population cible, et/ou lorsque les témoins ne sont pas issus de la même population que les cas. Dans cette étude, tous les cas correspondant à la population cible ont été sélectionnés, et les témoins tirés au sort dans la même population (l'ensemble des enfants nés en Norvège entre 1996 et 2001). Cependant, 88 % des cas identifiés, et 76 % des témoins tirés au sort ont accepté de participer à l'étude. Les mères ayant refusé de participer peuvent présenter des caractéristiques particulières, ce qui peut induire un biais de sélection. Par ailleurs les mères ne parlant pas le norvégien, de même de même que les mères d'un enfant décédé, ce qui peut également conduire à un biais. Une façon d'identifier ce biais et de donner des pistes sur son impact sur les résultats est de comparer, lorsque cela est possible, les caractéristiques des participants à celles de non participants.

2. Biais de confusion : dans les études observationnelles, ce biais est toujours possible.

3. Biais de classement : la consommation de café est mesurée à partir des déclarations des mères et de façon rétrospective ; ceci expose à un biais de classement vis-à-vis de cette exposition. A contrario, un biais de classement est peu probable concernant pour le classement en cas et témoins.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Cette étude n'est pas concernée par le biais d'attrition. Ce biais intervient dans les études longitudinales expérimentales ou non, dans lesquelles des sujets inclus peuvent sortir de l'étude en cours de suivi, et la population analysée ne correspond alors pas à la population incluse.

### Question 14

**Réponses exactes : B – D**

*Commentaires* Pour ce type de question, il est important de se remémorer les critères de causalité de Bradford Hill.

1. Réalité (impact limité des biais et des fluctuations d'échantillonnage) item A faux : les principaux facteurs de confusion sont pris en compte mais on ne peut exclure un effet de confusion résiduel ; l'ajustement réalisé ne suffit pas à argumenter une relation causale.

2. Force de l'association (mesurée par risque relatif ou Odds ratio ; plus les valeurs sont loin de 1, plus l'association est forte) ; ici les associations sont modérées.

3. Relation temporelle (temporalité) : les causes doivent précéder les conséquences ; c'est un élément nécessaire et donc essentiel à considérer : la temporalité dans une étude cas-témoin est difficile à établir ; on observe plutôt une coexistence qu'une antériorité.

4. Spécificité (une cause produit un effet et un seul et cet effet n'est présent que lorsque la cause est présente) ; il s'agit d'un critère rarement vérifié.

5. Relation dose-effet en faveur d'un effet directionnel (un changement dans la cause produit un changement dans l'effet) ; il s'agit d'un argument important de causalité vérifié ici par l'augmentation de l'OR en fonction de l'augmentation de la consommation de café et les tests de tendance significatifs.

6. Cohérence (répétition des observations dans différentes populations et par différents schémas d'étude).

7. Plausibilité (plausibilité biologique).

A. Preuve expérimentale (chez l'animal ou chez l'homme) ; un argument en faveur est développé dans l'introduction.

B. Analogie.

La proposition E n'est pas un argument en faveur de la nature causale de la relation entre café et fente labio-palatine, puisqu'elle concerne les fentes palatines seules dont l'étiologie pourrait être différente (évoqué dans la partie statistique).

### Question 15

**Réponses exactes : D – E**

*Commentaires* Il est bien inenvisageable sur le plan éthique (comme d'ailleurs sur le plan de la faisabilité) de réaliser une étude randomisée thé *versus* café (item A faux). Cette étude bien que de très bonne qualité, n'a pas le niveau de preuve suffisant (niveau 3) pour conduire à la recommandation d'une consommation de plus de 3 tasses de thé par jour (item B faux), ou d'une surveillance échographique renforcée génératrice d'une anxiété importante chez les futures mères (item D faux).

En revanche, les deux derniers items semblent raisonnables : en l'absence de preuve d'innocuité, il semble légitime de conseiller d'éviter leur consommation au cours du premier trimestre aux regard des résultats de cette étude. C'est un principe de précaution que l'on applique largement, notamment à tous les médicaments pour lesquels on ne dispose pas de donnée pendant la grossesse.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

# Dossier 2 Épreuve blanche nationale LCA 2017

Hervé Devilliers, Julie Berbis, Moustapha Dramé, Christine Binquet

| | |
|---|---|
| **Titre de l'article** | **Mortality results from the Götebord Randomised Prostate Cancer Screening Trial** |
| **Auteurs** | Hugosson J, Carlsson S, Aus G, Bergdhl S, Khatami A, Lodding P, Pihl CG, Stranne J, Holmberg E, Lija H |
| **Référence** | Lancet Oncol 2010 ; 11(8) :725-32 |
| **Lien vers l'article** | http://doi.org/10.1016/S1470-2045(10)70146-7 |

**1** **Les raisons invoquées par les auteurs pour justifier de la réalisation de cette étude sont que les études précédemment conduites**

a. Avaient été conduites de façon rétrospective
b. N'avaient pas pour certaines inclus suffisamment de patients
c. N'avaient pas permis de lever la controverse sur le rapport bénéfice/risque
d. N'étaient pas randomisées
e. N'avaient pas pris un temps de suivi suffisant pour juger d'un effet sur la mortalité

**2** **Les auteurs ont choisi de réaliser l'étude en population parce qu'ils souhaitaient privilégier**

a. La meilleure généralisation possible des résultats
b. La qualité de la randomisation en deux groupes parallèles
c. Les aspects éthiques liés à ce type d'étude
d. La procédure d'assignation masquée des patients dans chacun des groupes
e. Le contrôle d'un biais de classement

**3** **Le registre de population qui recense au 31 décembre 1994 les 32 298 hommes nés entre 1930 et 1944 à Göteborg peut être considéré comme**

a. La population source de l'étude
b. Un registre épidémiologique des cancers
c. L'échantillon de l'étude
d. La cohorte des patients
e. Le groupe témoin



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**4** **À propos de l'échantillon de 20 000 individus**

a. La procédure de tirage au sort s'apparente à en sondage en grappes
b. Chaque groupe comportait avant la randomisation le même nombre de personnes
c. La structure en âge de l'échantillon de 20 00 hommes était comparable à celle du registre populationnel
d. L'échantillon a été constitué par tirage au sort à partir du registre populationnel
e. La procédure de tirage au sort a été stratifiée en fonction de la cohorte de naissance

**5** **Il s'agit d'une étude**

a. Longitudinale
b. Étiologique
c. Cas-témoins
d. Pronostique
e. Interventionnelle

**6** **Les 2 objectifs principaux de cette étude interventionnelle randomisée et comparative étaient d'évaluer**

a. La proportion d'hommes présentant une élévation de PSA
b. La prévalence annuelle du cancer de la prostate chez les hommes de plus de 50 ans habitant la ville de Göteborg, Suède
c. L'efficacité clinique comparée d'une stratégie de dépistage du cancer de la prostate tous les deux ans *versus* la pratique courante
d. Le degré d'acceptation d'une stratégie de dépistage du cancer de la prostate tous les 2 ans
e. Les performances diagnostiques du dosage de PSA dans le diagnostic du cancer de la prostate

**7** **Les auteurs font référence au calcul de l'indicateur NNS (« number needed to be screened »). On peut dire de cet indicateur que :**

a. Il donne le nombre d'événements évités pour 100 patients dépistés
b. Il fait référence à la réduction relative de risque de décès par cancer de la prostate durant la période de suivi considérée
c. Il est calculé comme le NNT (number needed to treat) pour la mesure d'efficacité d'un essai thérapeutique
d. Il fait référence à la fraction étiologique du risque de décès
e. Il donne le nombre moyen de patients à dépister pour éviter un décès par cancer de la prostate durant la période de suivi considérée



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**8** **Dans cet article, un tableau de comparaison des groupes après randomisation n'a pas été fourni par les auteurs. Que peut-on en dire ?**

a. Le tableau de comparaison est sans intérêt lorsque la randomisation a été correctement effectuée
b. Ce tableau n'aurait servi à rien car tous les patients sont des hommes
c. Un tel tableau n'aurait servi à rien car tous les participants sont issus d'une même population
d. Le nombre élevé de patients randomisés rend un tel tableau inutile
e. La comparabilité de la structure d'âge de chaque groupe après randomisation peut être vérifiée dans tableau 4

**9** **À propos de la figure 2**

a. La médiane de survenue d'un cancer de la prostate dans le groupe dépistage (« screening ») est d'environ 6 ans
b. Elle est construite à partir des estimateurs de Kaplan-Meier
c. L'incidence cumulée de cancer de la prostate dans le groupe contrôle est d'environ 5 % à 10 ans
d. On vérifie aisément les données du texte relatives aux valeurs de l'incidence cumulée de cancer de la prostate après 14 ans de suivi
e. Il s'agit d'un histogramme des fréquences cumulées

**10** **La diminution des effectifs en regard des années 5, 10 et 14 de l'axe des abscisses de la figure 2 s'explique par**

a. Les perdus de vue
b. Les patients décédés d'autres causes
c. La survenue d'un cancer de la prostate
d. Les patients décédés d'un cancer de la prostate
e. Les patients définis comme « non attendees » ou « non participants »

**11** **Les auteurs on fait appel à un comité d'adjudication (« COD committee ») pour statuer sur la cause des décès en aveugle du groupe de randomisation. On peut en déduire que**

a. L'étude est en double aveugle
b. Le critère de jugement sur la mortalité liée au cancer de la prostate est un critère composite puisqu'il repose sur deux évaluations : l'avis du comité et la déclaration de décès
c. Le biais de suivi est minimisé
d. Le biais de classement ou d'évaluation sur le critère de jugement principal a été minimisé
e. Les investigateurs ont souhaité limiter les biais de confusion



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

**12** **Que peut-on dire du résultat concernant le risque relatif de décès par cancer de la prostate (RR estimé à 0,56, IC à 95 %, [0,39 à 0,82], p = 0,002) ?**

a. Le RR est statistiquement différent de 0 parce que $p < 0,05$
b. La comparaison des taux cumulés de mortalité par cancer de la prostate entre les deux groupes à une puissance de 99,8 %
c. Le RR est statistiquement significatif car compris entre les bornes de son intervalle de confiance à 95 %
d. Le risque d'erreur de 1re espèce est égal à 0,002
e. La réduction relative de risque de décès par cancer de la prostate est égale à 44 %

**13** **A l'analyse du flow chart, que peut-on déduire ?**

a. Les 48 patients exclus de chaque groupe après randomisation sont des sorties d'études.
b. On peut calculer la létalité du cancer de la prostate.
c. Le décès par cancer de la prostate sont des décès incidents.
d. Le diagramme se termine avec les effectifs de fin d'étude.
e. Il y a eu un appariement 1:1 entre les patients invités au dépistage et ceux du groupe contrôle.

**14** **Un patient de 80 ans, suite au décès de son meilleur ami de 75 ans d'un cancer de prostate, vous demande une prescription pour un dépistage par mesure du PSA. À la lecture de cet article, quels sont les arguments défaveur du dépistage du cancer de la prostate chez ce patient ?**

a. L'intérêt du dépistage pour le diagnostic de cancer de la prostate n'est pas démontré
b. La balance bénéfices/risques est défavorable après 70 ans
c. Le nombre de sujet nécessaire à dépister (NNS) est trop élevé pour rendre le dépistage systématique pertinent
d. En l'absence de symptôme, son risque de cancer de la prostate est négligeable.
e. Il y a un risque de surdosage et de sur traitement

**15** **Parmi les éléments suivants, quels sont ceux qui traduisent le respect par les auteurs des règles de transparence ou d'indépendance ?**

a. L'étude a été référencée dans un registre des essais cliniques sous le numéro ISRCTN54449243
b. L'étude est multicentrique
c. Le financeur n'a joué aucun rôle dans la conception et la réalisation de l'étude
d. Les auteurs ont publié les sources de financement de l'étude
e. Les auteurs ont publié les liens d'intérêts



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 1

**Réponses exactes : B – C – E**

*Commentaires* Cette étude était précédée d'au moins deux études randomisées (l'une européenne et l'autre américaine – item D faux). Ces études étaient par définition des études prospectives. Il n'est en effet pas possible de randomiser des patients dans différents groupes d'intervention *a posteriori*. L'item A est donc faux. En revanche, ces études pouvaient de taille insuffisante (item B vrai) ; d'autres études étaient par ailleurs critiquables sur le plan méthodologique. L'item C peut être considéré comme vrai puisque l'étude américaine n'apporte pas de résultat concluant quant à la mortalité liée au cancer de la prostate, contrairement à l'étude européenne qui retrouvait un résultat significatif sur ce critère. Par ailleurs ces études pouvaient être considérées comme ayant un suivi insuffisant. En effet, l'étude européenne bien que significative après un suivi de 9 ans aboutissait à un nombre de sujets à dépister pour prévenir un décès était très élevé du fait d'un nombre d'événements insuffisant (Item E vrai).

### Question 2

**Réponse exacte : A**

*Commentaires* Une étude menée en population est censée inclure une population moins sélectionnée qu'une étude menée dans un centre spécifique ; la population a plus de chance d'être représentative de la population cible. Cela optimise forcément la généralisation des résultats (Item A vrai). Ici les résultats pourront être extrapolés à l'ensemble des patients suédois de 50 à 65 ans. La qualité de la randomisation n'est pas influencée par le fait que l'étude soit menée en population générale ou sur une population plus spécifique (item B faux). Une étude menée en population n'est pas forcément non plus une étude plus éthique qu'une étude menée sur une population plus sélectionnée (item C faux) ; tout dépend de la balance bénéfice/risque pour les individus participant à l'essai. En plus ici l'étude est menée sans information préalable des sujets concernant l'essai, ce qui peut se justifier compte tenu de la question posée, mais n'est pas particulièrement classique. L'étude est réalisée en ouvert, l'item D est donc faux. Le biais de classement correspond à une erreur dans la mesure d'une exposition ou d'un critère de jugement. Ici l'exposition correspond au fait de recevoir ou non l'intervention, le critère de jugement est la mortalité. Une étude en population souvent de taille plus large qu'une étude classique nécessite souvent une organisation plus lourde pour arriver au même niveau de qualité des données. Le contrôle des biais de classement n'est donc pas plus simple.

### Question 3

**Réponse exacte : A**

*Commentaires* La population source est la population dont est tiré l'échantillon. Les hommes inscrits dans le registre de Göteborg constituent donc bien la population source (item A vrai). Il ne constitue pas l'échantillon d'étude (20 000 hommes tirés au sort – item C faux). Idéalement la population source doit être représentative de la population-cible (celle à laquelle on veut généraliser les résultats). L'échantillon est représentatif si tous les patients de la population-source ont la même probabilité de se trouver dans l'échantillon d'étude. D'autres exemples de populations sources pourraient être : « tous les patients se présentant dans les maternités de la région xx pour suspicion de menace d'accouchement prématurée » ou encore « tous les patients vus au CHU de xx pour une douleur thoracique ».



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Un registre de population n'est pas un registre de cancer : un registre de population inclut toutes les naissances, un registre de cancer n'inclut que les cas de cancers avérés (item B faux). L'item D est faux : le registre n'inclut pas des patients mais des hommes nés entre 1930 et 1944 qui peuvent être indemnes de cancer. Le registre ne correspond pas non plus au seul groupe témoin, puisque celui-ci correspond au seul groupe des hommes randomisés dans le groupe contrôle.

### Question 4

**Réponses exactes : C – D – E**

*Commentaires* L'item A est faux car une randomisation individuelle a été réalisée. Dans une randomisation en grappe, ou en cluster, l'unité de randomisation comporte plusieurs individus (par exemple, on tire au sort les écoles qui vont bénéficier d'un programme de prévention, les régions ou les hôpitaux dans lesquels on va mettre en place une nouvelle technique, etc.). Ici l'unité de randomisation est bien le patient.

Le tirage au sort a été réalisé de façon à ce que le nombre d'hommes dans chaque cohorte de naissance soit proportionnel au nombre dans la cohorte d'origine, ce qui correspond à un tirage au sort stratifié en fonction de la cohorte de naissance. La réponse B est fausse au contraire des réponses C et E. La réponse D est également vraie.

### Question 5

**Réponses exactes : A – E**

*Commentaires* Les patients sont inclus et bénéficient d'un suivi de 15 ans (item A vrai). La réponse B est fausse, le but ici n'est pas de documenter la cause d'une maladie. Par ailleurs le facteur d'exposition étudié ici est contrôlé et alloué de façon aléatoire. Ce schéma ne correspond pas à une étude étiologique dans laquelle on contrôlera éventuellement la maladie mais pas les facteurs d'exposition qui sont l'objet de la recherche. Une étude cas-témoin est généralement préférée dans ce cas, mais ce n'était pas le schéma d'étude ici (item C faux). L'item D est également faux : dans une étude pronostique, on inclut des patients atteints d'une maladie et on étudie les facteurs associés à la mortalité ou à l'évolution de cette maladie. Dans cette étude, la plupart des hommes inclus n'ont pas de cancer de la prostate. L'item E est en revanche vrai : dans le cadre de l'étude on propose aux hommes quelque chose (une intervention) en plus du suivi habituel.

### Question 6

**Réponses exactes : C – D**

*Commentaires* L'étude a pour objectif à la fois d'étudier l'acceptabilité du dépistage dans la population ainsi que son efficacité réelle (effectiveness par opposition à efficacy qui correspond à l'efficacité expérimentale). L'item A qui correspond à un objectif descriptif est donc faux, de même que l'item B. On s'intéresse à la détection de nouveaux cas de cancer et à la mortalité dans la population pas à la prévalence annuelle du cancer. L'item C est, en revanche vrai et fait bien référence au critère de jugement principal. L'item D est également vrai. L'item E est bien évidemment faux : il n'y a pas de calcul de sensibilité ou de spécificité et seuls les patients positifs bénéficient de l'examen urologique.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

### Question 7

**Réponses exactes : C – E**

*Commentaires* Le NNS est le nombre moyen de patients à dépister sur une durée donnée pour éviter un décès ou un effet secondaire (item E vrai). Dans cette étude, il s'agit du nombre moyen de patients à inviter au dépistage pour éviter un décès par cancer de prostate. C'est une statistique qui permet d'identifier, si elle est trop élevée, une stratégie inappropriée de dépistage de masse. L'item A est donc faux. L'item B est également faux car il s'agit de l'inverse de la réduction absolue d'après le chapitre matériel et méthode (par exemple s'il y a 0,5 % de décès par cancer dans un groupe et 0,30 % dans l'autre, la réduction absolue est de 0,5-0,3 = 0,2 % et le NNS est de 1/0,002 = 500 patients à dépister en moyenne pour éviter un cancer). L'item C est vrai : dans un essai clinique le NNT pour éviter un évènement (un décès de cause cardiovasculaire sous statine, une fracture sous biphosphonates, etc. est 1/(fréquence sans traitement – fréquence avec traitement). L'item D est faux ; le NNS ne fait pas référence à la FER, qui est la proportion de cas attribuables (ici décès) à l'exposition (ici plutôt absence d'exposition) parmi les sujets exposés. Le calcul s'appuie sur la mesure d'association (RR ou OR).

### Question 8

**Réponse exacte : E**

*Commentaires* L'item A est faux : c'est justement pour vérifier la comparabilité des groupes que l'on a besoin de ce tableau. Idem pour l'item B : même s'il n'y a que des hommes, il faut vérifier que les autres caractéristiques sont comparables entre les deux groupes. Quand bien même les sujets sont issus d'une même population, il faut encore vérifier que la randomisation a permis de les répartir en deux groupes aux caractéristiques comparables par la randomisation (item C faux). Le nombre élevé de sujets ne fait rien à l'affaire, le tableau est indispensable (item D faux). En revanche l'item E est vrai.

### Question 9

**Réponses exactes : C – D**

*Commentaires* La médiane de survenue d'un cancer est le délai au-delà duquel la moitié des patients ont présenté un cancer. La médiane de survenue d'un cancer serait donc la durée correspondante à une incidence cumulée de 50 %. D'après la figure 2, l'incidence cumulée reste largement inférieure à 50 % même à 14 ans. L'item A est donc faux. L'item B par contre est vrai (*cf.* partie statistiques). L'item C est également vrai : il suffit de vérifier sur la courbe la durée correspondante à 5 % d'incidence cumulée. Pour l'item D (item vrai) le texte indique « The cumulative incidence of prostate cancer at 14 years was 12,7 % in the screening group *versus* 8,2 % in the control group, corresponding to a hazard ratio of 1,64 (95 % confidence interval [CI]= 1,50-1,80, p < 0,0001). L'item E est en revanche faux : un histogramme est une représentation graphique de la distribution d'une variable quantitative (ici ce pourrait être le délai d'apparition d'un cancer chez les sujets concernés, alors que la figure 2 représente la probabilité cumulée de survenue d'un cancer de la prostate dans la population étudiée au cours du temps.

### Question 10

**Réponses exactes : A – B – C**

*Commentaires* Les patients « at risk » sont les patients encore suivis (donc non décédés ou émigrés), qui n'ont pas encore présenté l'évènement. Cette étude s'appuie sur un registre de cancer (identification exhaustive des cas de cancer et sur l'état civil, la notion de perdue de vue ne s'applique pas bien ici (item A contestable). Le seul cas d'arrêt du suivi est lié à une émigra-



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

tion des hommes. On parlera donc plutôt d'une censure ; dans ce cas elle est aléatoire car non liée au critère de jugement. Les décès d'une autre cause sont dans le même cas ; il s'agit aussi d'une censure qui peut être considérée comme compétitive par rapport au risque de cancer de la prostate. Les sujets décédés d'une autre cause avant le diagnostic d'un cancer de la prostate n'ont pas eu le temps de développer un tel cancer. Les patients atteints d'un cancer de la prostate ne contribuent plus aux effectifs « at risk » une fois celui-ci diagnostiqué. Ce phénomène participe bien à la diminution des effectifs mais parce qu'ils ont déjà eu l'événement d'intérêt. Il s'agit donc d'un cas différent des réponses A et B. Les patients décédés d'un cancer de la prostate ne contribuaient plus aux effectifs à risque dès le diagnostic de celui-ci. Seuls les cas de diagnostic de cancer de la prostate au moment du décès pourraient contribuer à la diminution des effectifs. Cette situation est certainement moins fréquente dans le groupe contrôle que dans le groupe test. L'item E est faux, car les modalités d'identification des cas de cancer sont indépendantes de la participation des sujets.

### Question 11

**Réponse exacte : D**

*Commentaires* Cette étude est par définition réalisée en ouvert (item A faux). La mise en place d'un comité d'adjudication permettant de classer les événements présentés par les sujets est habituellement une mesure qui permet d'éviter le biais de mesure (=biais de classement) sur le critère de jugement principal (item D vrai). L'item B est faux : ce n'est pas la définition d'un critère composite. Un critère composite est un critère définit par la survenue d'au moins un événement parmi les événements composant ce critère. C'est le cas par exemple de la survie sans progression, pour laquelle on considère que le patient a l'événement d'intérêt s'il progresse ou s'il décède. Un autre critère composite classique est la survenue d'un évènement cardiovasculaire majeur (il peut s'agir d'un IDM ou d'un AVC, fatal ou non). L'item C est également faux : le biais de suivi est lié au fait que les deux groupes ne sont pas suivi de la même manière. Ceci est essentiellement lié à l'absence d'aveugle. L'item E est faux : la randomisation contribue à limiter les biais de confusion mais pas les comités d'adjudication.

### Question 12

**Réponse exacte : E**

*Commentaires* Un RR s'interprète en cherchant à savoir s'il est significativement différent 1, puisqu'il s'agit d'un rapport d'incidence et qu'on cherche à savoir s'il y a une réelle différence d'incidence entre les groupes (s'il n'y a pas de différence, numérateur = dénominateur donc forcément RR = 1). Et donc cela n'a pas de sens de comparer le RR à la valeur 0 (item A faux). Pour ce qui concerne l'item B, 99,8 % correspond au complémentaire de la p-value ici égale à 2 pour 1 000. Il s'agit donc de la probabilité que la différence observée ne soit pas liée au hasard. Cela est complètement différent de la puissance qui correspond à la probabilité de mettre en évidence la différence si elle existe. Pour rappel, ci-dessous le tableau résumé des différents types d'erreurs.

**Tableau 1. Types d'erreurs et test d'hypothèses**

| Décision | Réalité | |
|---|---|---|
| | $H_0$ est vraie | $H_1$ est vraie |
| Ne pas rejeter $H_0$ | $1-\alpha$ | $\beta$<br>Erreur de deuxième espèce |
| Rejeter $H_0$ | $\alpha$<br>Erreur de première espèce | $1-\beta$<br>Puissance |



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni

Pour ce qui concerne l'item C, par définition une estimation d'une mesure d'association doit être contenue dans les bornes de son intervalle de confiance. Si ce n'est pas le cas, c'est qu'il y a une coquille d'impression. En tout cas cela ne donne aucune indication sur la significativité de l'association (item C faux). Le risque d'erreur de 1re espèce est le risque de rejeter l'hypothèse nulle° alors que celle-ci est vraie. C'est un risque fixé *a priori* (en général 0,05 en formulation bilatéral). La p-value est la probabilité, en supposant que Ho est vraie, d'obtenir une valeur de la variable de décision (ici RR) au moins aussi éloignée de 1 que la valeur du RR que l'on a obtenue avec notre échantillon. En gros cela revient à la probabilité d'observer un RR d'au moins 0,82 avec notre échantillon alors que le RR = 1 en réalité. Le risque de 1re espèce et la p-value sont donc différents, même si le 1er permet d'interpréter le second et l'item D est faux. Par contre l'item E est vrai : le RR est de 0,56, ce qui correspond à une réduction du risque relatif de 44 % (1-0,56)

**Question 13**

**Réponse exacte : A**

*Commentaires* Les patients exclus avant la randomisation ne sont pas des sortis d'étude. Les cancers prévalents ne répondent pas aux critères de sélection de l'étude, il s'agit de patients « inclus à tort » (item A faux). Dans cette étude, on peut calculer la proportion de décès par rapport au nombre de cas atteints d'un cancer de la prostate dans l'échantillon d'étude, ce qui correspond à la définition du taux de létalité. Cependant, le taux de létalité doit être calculé dans une population représentative des patients atteints d'un cancer de la prostate suivi dans le cadre d'une prise en charge habituelle. Ici les patients diagnostiqués dans le cadre du dépistage sont diagnostiqués à un stade plus précoce et donc avec un risque plus limité de décès, ce qui conduit à modifier l'estimation globale. L'item B est donc faux. L'item C est vrai : autant on peut parler de cancer prévalent et compter l'ensemble des patients atteints de cancer dans une population quelle que soit leur ancienneté, autant la notion de décès, ne peut être envisagée sous cet angle : le concept de décès prévalent n'existe pas. Il s'agit donc forcément des nouveaux cas de décès et donc de cancers incidents. L'item D est faux, les effectifs correspondent aux événements d'intérêt (cancer de la prostate et décès lié au cancer. Enfin, le terme d'appariement n'est pas adapté à cette étude. Il s'agit d'une allocation par randomisation.

**Question 14**

**Réponse exacte : E**

*Commentaires* L'item A est faux car il y a bien une diminution de la mortalité dans le groupe dépistage par rapport au groupe contrôle. Si les groupes sont constitués en les stratifiant sur l'âge, l'article ne permet pas de le démontrer spécifiquement. Par contre le NNS est bien plus faible que ce qui est connu dans le cancer du sein ou du colon, l'item C est faux. L'item D est également faux : le cancer est souvent asymptomatique. L'item E est vrai.

**Question 15**

**Réponses exactes : A – C – D – E**

*Commentaires* L'item B est faux ; le caractère multicentrique est un argument favorisant la généralisation des résultats mais n'est pas un élément traduisant le respect par les auteurs des règles de transparence ou d'indépendance.



Originellement diffusé par Faille ECNi, groupe non lucratif de partage de connaissances médicales à tous. https://t.me/faillecni